



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 177427

TO: Christian Fronda  
Location: rem/2D78/2C70  
Art Unit: 1652  
Monday, January 30, 2006  
Case Serial Number: 10/663033

From: Barb O'Bryen  
Location: Biotech-Chem Library  
Remsen 1a69  
Phone: 571-272-2518

*BOB*  
barbara.obryen@uspto.gov

### Search Notes

*Sequences created & searched:*

*Seq 3 fused to reverse complement of Seq 4  
(3 - revcomp - 4)*

*Seq 4 reverse complement fused to Seq 3  
(revcomp 4 - seq 3)*

*gap size penalty changed from 0.5 (default)  
to 0.1*

**This Page Blank (uspto)**

STIC-Biotech/ChemLib

From: Fronda, Christian  
Sent: Tuesday, January 24, 2006 11:35 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search for Serial No. 10/663,033

Importance: High

For Serial No. 10/663,033, there is a claim directed toward a polynucleotide obtained by PCR using PCR primers of SEQ ID NO: 3 and SEQ ID NO: 4.

Is there a way to search for polynucleotides that are obtained by PCR primers in the commercial, issued, PGPub, and pending databases? I only need results that have both SEQ ID NO: 3 and the reverse complement of SEQ ID NO: 4.

Please advise. Thank you.

Christian L. Fronda  
Art Unit 1652  
Office REM 2D78  
Mailbox REM 2C70  
(571)272-0929

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: 1/30  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

This Page Blank (uspto)



GenCore version 5.1.6

Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:30:51 ; Search time 894 Seconds  
(without alignments)  
2479.748 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4

Perfect score: 39  
Sequence: 1 gaacactcacacgcgtttatagcagcgtctaattcta 39

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 5883141 seqs, 2842172553 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_at.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vl.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24.2	62.1	325543	14 AC152657	AC152657 Bos tauru
2	22.6	57.9	208312	14 CR352328	CR352328 Dabio rer
3	22	56.4	805	1 AF484440	AF484440 Helicobac
4	22	56.4	819	1 AF484441	AF484441 Helicobac
5	22	56.4	822	1 AF484437	AF484437 Helicobac
6	22	56.4	934	6 BD092531	BD092531 Identific
7	22	56.4	1037	1 AF093831	AF093831 Helicobac
8	22	56.4	1119	1 AF076779	AF076779 Helicobac
9	22	56.4	1119	6 AR154400	AR154400 Sequence
10	22	56.4	1119	6 BD238089	BD238089 Helicobac
11	22	56.4	1119	6 AR442784	AR442784 Sequence
12	22	56.4	1122	1 AF093828	AF093828 Helicobac
13	22	56.4	1124	1 AF093829	AF093829 Helicobac
14	22	56.4	1124	1 AF093832	AF093832 Helicobac
15	22	56.4	1124	1 AF093833	AF093833 Helicobac
16	22	56.4	1130	1 AF093830	AF093830 Helicobac
17	22	56.4	10530	1 AE000531	AE000531 Helicobac
18	22	56.4	166518	8 AC090514	AC090514 Homo sapi

c 19	21.6	55.4	3089	5	AJ719964	AJ719964 Gallus ga
c 20	21.6	55.4	102146	14	AC133396	AC133396 Felis cat
c 21	21.6	55.4	249683	14	AC119541	AC119541 Rattus no
c 22	21.4	54.9	110000	14	AP006493	Continuation (5 of
c 23	21.2	54.4	734	8	HS4333671	AJ333671 Homo sapi
c 24	21.2	54.4	101904	5	AL935310	AL935310 Zebrafish
c 25	21.2	54.4	152898	8	AL445240	AL445240 Human DNA
c 26	21.2	54.4	160737	14	AC026213	AC026213 Homo sapi
c 27	21.2	54.4	169675	14	AC146030	AC146030 Pan trogl
c 28	21.2	54.4	169856	14	AC144890	AC144890 Sus scrof
c 29	21.2	54.4	178009	14	AC119549	AC119549 Rattus no
c 30	21.2	54.4	220420	5	BX000452	BX000452 Zebrafish
c 31	21.2	54.4	318662	14	AC109547	AC109547 Rattus no
c 32	21	53.8	754	10	BV619750	BV619750 S216P6228
c 33	21	53.8	38688	14	AC100280	AC100280 Mus muscu
c 34	21	53.8	77087	8	AL355877	AL355877 Human DNA
c 35	21	53.8	80279	8	HS75K24	AL035700 Human DNA
c 36	21	53.8	150311	9	AC112081	AC112081 Rattus no
c 37	21	53.8	163500	14	AL450316	AL450316 Homo sapi
c 38	21	53.8	175172	14	AL355859	AL355859 Homo sapi
c 39	21	53.8	176135	9	AC161165	AC161165 Mus muscu
c 40	20.8	53.3	335	8	HS191XH12	Z23462 H. sapiens
c 41	20.8	53.3	95546	8	AC022816	AC022816 Homo sapi
c 42	20.6	52.8	373	10	BV094898	BV094898 RPAMMSQ0
c 43	20.6	52.8	373	10	BV100363	BV100363 RPAMMSQ0
c 44	20.6	52.8	373	10	BV156913	BV156913 RPAMMSQ0
c 45	20.6	52.8	9033	7	BPAJ7641	AJ000741 Bacteriop

## ALIGNMENTS

## RESULT 1

## AC152657

## LOCUS

## DEFINITION

Bos taurus clone CH240-10013, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 46

unordered pieces.

## AC152657

## AC152657.4

GI:68226917

HTG: HTGS PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.

Bos taurus (cow)

## SOURCE

Bos taurus

## REFERENCES

1 (bases 1 to 325543)

Muzny D.Marie., Metzker M.Lee., Abramzon S., Adams C., Alder J.,

Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,

Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,

Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,

Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,

Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,

Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,

Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,

Cleveland L., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,

Davila M.C., Davis C., Davy-Carroll L., De Anda C., Dederich D.,

Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K.,

Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,

Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,

Fernandes S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,

Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,

Gebregorgis E., Geer K., Gill R., Grady M., Guerra W., Guevara W.,

Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,

Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,

Hernandez R., Hines S., Hladun S.L., Hodgson A., Hoggess M.,

Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,

Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,

Karpachy S., Kelly S., Khan Z., King L., Kovar C.,

Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,

Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,

Lorensuewa L., Loulseghe H., Lozano R.J., Lu X., Ma J.,

Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,

Mangum B., Mapua P., Martin K., Martin R., Martinez E.,

AC152657 Bos taurus clone CH240-10013, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 46

AC152657

AC152657.4

GI:68226917

HTG: HTGS PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.

Bos taurus (cow)

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Bos taurus

Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwakoelameh, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C., Plopper, F., Poldinger, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 325543)  
Worley, K.C.  
Direct Submission  
Submitted (16-NOV-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 325543)  
Cow Genome Sequencing Consortium.  
Direct Submission  
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: FAOI  
Center clone name: CH240-10013  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 304082 bases at least Q40  
Consensus quality: 307999 bases at least Q30  
Consensus quality: 311731 bases at least Q20  
Estimated insert size: 312171; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html)  
\* NOTE: This sequence may represent more than one 'clone'  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 46 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
1647: contig of 1647 bp in length  
1697: gap of 50 bp  
6641: contig of 4944 bp in length  
6691: gap of 50 bp  
6692: contig of 13137 bp in length  
19829: gap of 50 bp  
19878: gap of 50 bp  
22273: contig of 2395 bp in length  
22773: gap of unknown length  
22774: contig of 6299 bp in length  
28672: gap of 50 bp  
28673: contig of 2106 bp in length  
30829: gap of unknown length  
30929: contig of 2277 bp in length  
33206: gap of 718 bp  
33924: contig of 3541 bp in length  
37465: gap of unknown length  
37465: contig of 9001 bp in length  
46566: gap of 50 bp  
46566: contig of 7570 bp in length  
54185: gap of 322 bp  
54507: gap of 518 bp  
75815: contig of 21307 bp in length  
76332: gap of 1592 bp in length  
77925: gap of 285 bp  
78209: contig of 3770 bp in length  
81979: gap of 50 bp  
82029: contig of 1693 bp in length  
83722: gap of 50 bp  
83723: contig of 13156 bp in length  
96928: gap of 50 bp  
96929: contig of 8387 bp in length  
103365: gap of 50 bp  
105415: contig of 7870 bp in length  
113285: gap of 232 bp  
113286: contig of 20597 bp in length  
134118: gap of 1574 bp  
135688: gap of 6916 bp in length  
142604: gap of 50 bp  
142605: contig of 15062 bp in length  
157716: gap of 50 bp  
157766: contig of 8847 bp in length  
166613: gap of unknown length  
166713: contig of 4303 bp in length  
17016: gap of 50 bp  
171066: contig of 4850 bp in length  
175916: gap of 50 bp  
175966: contig of 3037 bp in length  
179003: gap of 632 bp  
179635: contig of 1467 bp in length  
181102: gap of 1130 bp  
182232: contig of 5664 bp in length  
182233: gap of 390 bp  
188286: contig of 9340 bp in length  
197626: gap of 50 bp  
197676: contig of 9699 bp in length  
207375: gap of 877 bp  
208252: contig of 7746 bp in length  
215998: gap of 50 bp  
216048: contig of 20173 bp in length  
236221: gap of 50 bp  
236222: contig of 17879 bp in length  
254150: gap of 50 bp  
254200: contig of 7291 bp in length  
261491: gap of 50 bp  
261492: contig of 14204 bp in length  
275745: gap of unknown length  
275845: contig of 4059 bp in length  
279904: gap of 50 bp  
279954: contig of 3432 bp in length  
283386: gap of 50 bp  
283387

```

* 283437 292331: contig of 895 bp in length
* 292332 292381: gap of 50 bp
* 292382 295701: contig of 3320 bp in length
* 295702 296854: gap of 1153 bp
* 296855 305830: contig of 8976 bp in length
* 305831 306127: gap of 297 bp
* 306128 310233: contig of 4106 bp in length
* 310234 310285: gap of 52 bp
* 310286 317330: contig of 7045 bp in length
* 317331 317430: gap of unknown length
* 317431 318430: contig of 1000 bp in length
* 318431 318530: gap of unknown length
* 318531 319535: contig of 1005 bp in length
* 319536 319635: gap of unknown length
* 319636 320712: contig of 1077 bp in length

Query Match      62.1%; Score 24.2; DB 14; Length 325543;
Best Local Similarity 78.4%; Pred. No. 3.8;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 ACACCTCACGCGTCTTATAGCAGCGTCTAATTTCTA 39
    |||||
Db 18261 AGACTCATACATCTTCTAGCTATGCTGCTAATTTTA 18297

RESULT 2
CR352328/c
LOCUS      208312 bp      DNA      linear      HTG 05-MAR-2005
DEFINITION Danio rerio clone CH211-212K18, WORKING DRAFT SEQUENCE, 3 unordered
            pieces.
ACCESSION  CR352328
VERSION    CR352328.5 GI:60543659
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Danio rerio (zebrafish)
            Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
            McLaren, S.
            Direct Submission
            Submitted (04-MAR-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Mar 5, 2005 this sequence version replaced gi:46981214.
            ----- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfsh-help@sanger.ac.uk
            ----- Project Information
            Center project name: zC212K18
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 207646 bases at least Q40
            Consensus quality: 207804 bases at least Q30
            Consensus quality: 207867 bases at least Q20
            Insert size: 208112; sum-of-contigs
            Insert size: 206965; 7.7% error; agarose-fp
            Quality coverage: 10.35x in Q20 bases; sum-of-contigs Quality
            coverage: 10.53x in Q20 bases; agarose-fp
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * 1 101194: contig of 101194 bp in length
            * 101195 101294: gap of 100 bp

```

```

* 101295 192622: contig of 91328 bp in length
* 192623 192722: gap of 100 bp
* 192723 208312: contig of 15590 bp in length.

FEATURES             source
     source
     1..208312
         /organism="Danio rerio"
         /mol_type="genomic DNA"
         /db_xref="taxon:7955"
         /clone="CH211-212K18"
         /clone_lib="CHORI-211"
     misc_feature
     1..101194
         /note="assembly fragment:01920
         fragment chain:1
         clone_end:8f6
         vector_end:8f6
         101295..192622
         /note="assembly fragment:00203
         fragment chain:1
         clone_end:17
         vector_end:right"
     misc_feature
     192723..208312
         /note="assembly fragment:00017"

ORIGIN

Query Match      57.9%; Score 22.6; DB 14; Length 208312;
Best Local Similarity 75.7%; Pred. No. 22;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAACACTCACGCGTCTTATAGCAGCGTCTAATTC 37
    |||||
Db 20088 GAACACTCCAGGTTTCTACAGCGAGGTCTACTTC 20052

RESULT 3
AF484440/c
LOCUS      805 bp      DNA      linear      BCT 05-MAY-2005
DEFINITION Helicobacter pylori isolate 60190 HpyIIIM protein (hpyIIIM) gene,
            partial cds.
ACCESSION  AF484440
VERSION    AF484440.1 GI:33337332
KEYWORDS  Helicobacter pylori
            Helicobacter pylori
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Helicobacteraceae; Helicobacter.
            1 (bases 1 to 805)
            Ando, T., Blaser, M.J. and Wassenaar, T.M.
            Evolutionary history of hrgA, a gene replacing the restriction gene
            in an R-M gene locus of H. pylori
            Unpublished
            2 (bases 1 to 805)
            Ando, T., Blaser, M.J. and Wassenaar, T.M.
            Direct Submission
            Submitted (19-FEB-2002) First Department of Internal Medicine,
            Nagoya University School of Medicine, 65 Tsurumai-cho Showa-Ku,
            Nagoya, Aichi 466-8550, Japan
            Nagoya, Aichi 466-8550, Japan
            Location/Qualifiers
            1..805
                /organism="Helicobacter pylori"
                /mol_type="genomic DNA"
                /isolate="60190"
                /db_xref="taxon:210"
            1..>805
                /gene="hpyIIIM"
            1..>805
                /gene="hpyIIIM"
                /codon_start=1
                /transl_table=11
                /product="HpyIIIM protein"
                /protein_id="AAQ13334.1"
                /db_xref="GI:33337333"

FEATURES             source
     source
     1..805
         /organism="Helicobacter pylori"
         /mol_type="genomic DNA"
         /isolate="60190"
         /db_xref="taxon:210"
     1..>805
         /gene="hpyIIIM"
     1..>805
         /gene="hpyIIIM"
         /codon_start=1
         /transl_table=11
         /product="HpyIIIM protein"
         /protein_id="AAQ13334.1"
         /db_xref="GI:33337333"

translation="MKASCLSRPIKSDQLFTLYQGDCNEVLPOFENAFDLIFADPP
YFLSNDGLSIQSGKIVSYNKGWDKENGINDIFNCQMINNAKALKNYGLLSGT

```

YHNIFSLGRILQKLDKILNLTWQKTNPPNPFNSCRYLTHSAEQI IWARKSRKHVFP  
 NYEVLKKNNDKQMRDVPAPAIAPWEKANGKHTOKPLALLVLLMLASDGNLSLIGD  
 PFSGSTTGTIAANLLKQPGIEKESEFIKISMRKLELDARYKIKRSKI"

## ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 18 TTATAGCGAGCGTCTAATTCTA 39  
 Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 4  
 AF484441/c  
 LOCUS  
 DEFINITION Helicobacter pylori isolate 9627 HpyIIIM protein (hpyIIIM) linear BCT 05-MAY-2005  
 partial cds.

ACCESSION AF484441  
 VERSION AF484441.1 GI:33337334  
 KEYWORDS  
 SOURCE Helicobacter pylori  
 ORGANISM Helicobacter pylori  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 Helicobacteraceae; Helicobacter.

REFERENCE 1 (bases 1 to 819)  
 Ando, T., Blaser, M.J. and Wassenaar, T.M.  
 Evolutionary history of hrgA, a gene replacing the restriction gene  
 in an R-M gene locus of H. pylori

JOURNAL Unpublished  
 AUTHORS  
 REFERENCE 2 (bases 1 to 819)  
 Ando, T., Blaser, M.J. and Wassenaar, T.M.  
 Direct Submission  
 TITLE Submitted (19-FEB-2002) First Department of Internal Medicine,  
 Nagoya University School of Medicine, 65 Teurumai-cho Showa-ku,  
 Nagoya, Aichi 466-8550, Japan

FEATURES  
 source  
 1..819  
 /organism="Helicobacter pylori"  
 /mol\_type="genomic DNA"  
 /isolate="9627"  
 /db\_xref="taxon:210"

gene  
 1..819  
 /gene="hpyIIIM"

CDS  
 1..819  
 /codon\_start=1  
 /transl\_table=11  
 /product="HpyIIIM protein"  
 /protein\_id="AAQ13335.1"  
 /db\_xref="GI:33337335"  
 /translation="MKTPCLSRPIFKSQDQLFTLYQGDCNEVLQFENRPLIFADPPP  
 YFLSNDGLSIQSGKIVSVNGKDWKEGVNGIDFNYQWNNAKKALNTGSLISGT  
 YHNIFSLGRILQKLDKILNLTWQKTNPPNPFNSCRYLTHSAEQI IWARKSRKHVFP  
 NYEVLKKNNDKQMRDVPAPAIAPWEKANGKHTOKPLALLVLLMLASDGNLSLIGD  
 PFSGSTTGTIAANLLKREFIGIEKESEFIKISMRKLELDARYKIKRSKIDLNH"

## ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 819;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39  
 Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 5  
 AF484437/c  
 LOCUS  
 DEFINITION Helicobacter pylori isolate JP2 HpyIIIM protein (hpyIIIM) linear BCT 05-MAY-2005  
 partial cds.

ACCESSION AF484437  
 VERSION AF484437.1 GI:33337326  
 KEYWORDS  
 SOURCE Helicobacter pylori

ORGANISM Helicobacter pylori  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 Helicobacteraceae; Helicobacter.

REFERENCE 1 (bases 1 to 822)  
 Ando, T., Blaser, M.J. and Wassenaar, T.M.  
 Evolutionary history of hrgA, a gene replacing the restriction gene  
 in an R-M gene locus of H. pylori

JOURNAL Unpublished  
 AUTHORS  
 REFERENCE 2 (bases 1 to 822)  
 Ando, T., Blaser, M.J. and Wassenaar, T.M.  
 Direct Submission  
 TITLE Submitted (19-FEB-2002) First Department of Internal Medicine,  
 Nagoya University School of Medicine, 65 Teurumai-cho Showa-ku,  
 Nagoya, Aichi 466-8550, Japan

FEATURES  
 source  
 1..822  
 /organism="Helicobacter pylori"  
 /mol\_type="genomic DNA"  
 /isolate="JP2"  
 /db\_xref="taxon:210"

gene  
 1..822  
 /gene="hpyIIIM"

CDS  
 1..822  
 /gene="hpyIIIM"  
 /codon\_start=1  
 /transl\_table=11  
 /product="HpyIIIM protein"  
 /protein\_id="AAQ13331.1"  
 /db\_xref="GI:33337327"

/translation="MEVPLSRPIFKSQDQLFTLYQGDCNEVLQFENRPLIFADPPP  
 YFLSNDGLSIQSGKIVSVNGKDWKEGVNDIDFNYQWNNAKKALNTGSLISGT  
 YHNIFSLGRILQKLDKILNLTWQKTNPPNPFNSCRYLTHSAEQI IWARKSRKHVFP  
 NYEVLKKNNDKQMRDVPAPAIAPWEKANGKHTOKPLALLVLLMLASDGNLSLIGD  
 PFSGSTTGTIAANLLKQPGIEKESEFIKISMRKLELDARYKIKRSKIDLNK"

## ORIGIN

Query Match 56.4%; Score 22; DB 1; Length 822;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39  
 Db 785 TTATAGCGAGCGTCTAATTCTA 764

RESULT 6  
 BD092531/c

LOCUS  
 DEFINITION Identification of polynucleotides encoding novel helicobacter  
 polypeptides in the helicobacter genome.

ACCESSION BD092531  
 VERSION BD092531.1 GI:22638142  
 KEYWORDS JP 2001527393-A/312.  
 SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.  
 1 (bases 1 to 934)  
 Kleanthous, H., Garawi, A.A., Miller, C., Tomb, J.P. and Oomen, R.P.

REFERENCE 1 (bases 1 to 934)  
 Kleanthous, H., Garawi, A.A., Miller, C., Tomb, J.P. and Oomen, R.P.  
 Identification of polynucleotides encoding novel helicobacter  
 polypeptides in the helicobacter genome

JOURNAL Patent: JP 2001527393-A 312 25-DEC-2001;  
 MERIEUX ORAVAX SOCIETE EN NOM COLLECTIF PASTEUR MERIEUX SERUMS ET  
 VACCINS AGROBIOLOGICAL RESOURCES MINISTRY O, SA, HUMAN GENOME  
 SCIENCES INC

COMMENT  
 PN JP 2001527393-A/312  
 PD 25-DEC-2001  
 PF 01-APR-1998 JP 1998541947  
 PR 01-APR-1997 US 08/833457, 24-JUN-1997 US 08/881227 PR  
 29-JUL-1997 US 08/902615

PI	HAROLD KLEANTHOS, AMAL AL GARAWI, CHARLES MILLER, JEAN FRANCOIS TOMB.
PI	RAYMOND PETER OOMEN
PC	A01N43/04, A61K31/70
CC	Strandedness: Single;
CC	Topology: Linear;
PH	Key Location/Qualifiers.
FEATURES	Location/Qualifiers
source	1. .934
ORIGIN	
Query Match	56.4%; Score 22; DB 6; Length 934;
Best Local Similarity	100.0%; Pred. No. 53;
Matches	22; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	18 TTATAGCGAGCGTCTAATTCTA 39
Db	844 TTATAGCGAGCGTCTAATTCTA 823
RESULT 7	
AF093831	
LOCUS	1037 bp DNA linear BCT 13-MAY-1999
DEFINITION	Helicobacter pylori strain UA1218 alpha-1,2-fucosyltransferase gene, complete cds.
ACCESSION	AF093831
VERSION	AF093831.1 GI:4808594
KEYWORDS	
SOURCE	Helicobacter pylori
ORGANISM	Helicobacter pylori
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales
AUTHORS	Helicobacteraceae; Helicobacter.
TITLE	1 (bases 1 to 1037)
JOURNAL	Wang, G., Rasiko, D.A., Sherburne, R. and Taylor, D.E.
PUBMED	Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori: analysis of the alpha (1,2) fucosyltransferase gene
AUTHORS	Mol. Microbiol. 31 (4), 1265-1274 (1999)
AUTHORS	10096092
TITLE	2 (bases 1 to 1037)
JOURNAL	Wang, G. and Taylor, D.E.
PUBMED	Submitted (24-SEP-1998) Medical Microbiology and Immunology, University of Alberta, 1-51 Medical Science Building, Edmonton, AB T6G 2H7, Canada
FEATURES	Location/Qualifiers
source	1. .1037
ORIGIN	
Query Match	56.4%; Score 22; DB 1; Length 1037;
Best Local Similarity	100.0%; Pred. No. 53;
Matches	22; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	18 TTATAGCGAGCGTCTAATTCTA 39
Db	844 TTATAGCGAGCGTCTAATTCTA 823

```
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1119)
TITLE Taylor,D., Wang,G. and Palcic,M.
JOURNAL .alpha.1,2 fucosyltransferase
FEATURES Patent: US 6238894-A 1 29-MAY-2001;
source Location/Qualifiers
1..1119
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 56.4%; Score 22; DB 6; Length 1119;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTTAATTCTA 39
|||||
Db 1098 TTATAGCGAGCGTCTTAATTCTA 1119

RESULT 10
BD238089 1119 bp DNA linear PAT 17-JUL-2003
LOCUS Helicobacter pylori-origin alpha-1,2-fucosyltransferase.
DEFINITION BD238089
ACCESSION BD238089
VERSION BD238089.1 GI:33047859
KEYWORDS JP 2002528122-A/1.
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 1119)
AUTHORS Taylor,D.E., Wang,G. and Palcic,M.
TITLE Helicobacter pylori-origin alpha-1,2-fucosyltransferase
JOURNAL Patent: JP 2002528122-A 1 03-SEP-2002;
GOVERNORS OF THE UNIVERSITY OF ALBERTA
COMMENT OS Helicobacter pylori
PN JP 2002528122-A/1
PD 03-SEP-2002
PF 03-NOV-1999 JP 2000579755
PR 04-NOV-1998 US 60/107268,02-NOV-1999 US 09/433598 PI
DIANE E TAYLOR,GE WANG,MONICA PALCIC
PC C12N15/09,C07K16/40,C12N1/15,C12N1/19,C12N1/21,C12N5/10 PC
,C12N9/10,C12P19/18,
PC C12P21/08,C12Q1/48,C12Q1/68,G01N33/569/(C12N9/10,C12R1:01),
PC (C12N9/10,C12R1:19),(C12P19/18,C12R1:19),C12N15/00,C12N5/00 CC
Helicobacter pylori-origin alpha-1,2-fucosyltransferase FH Key
Location/Qualifiers
(137)..(1036).

FEATURES FT CDS
source Location/Qualifiers
1..1119
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/db_xref="taxon:210"

ORIGIN
Query Match 56.4%; Score 22; DB 6; Length 1119;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTTAATTCTA 39
|||||
Db 1098 TTATAGCGAGCGTCTTAATTCTA 1119

RESULT 11
AR442784 1119 bp DNA linear PAT 20-FEB-2004
LOCUS Sequence 1 from patent US 6670160.
DEFINITION AR442784
ACCESSION AR442784
VERSION AR442784.1 GI:42670204
KEYWORDS .
```

```
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1119)
TITLE Taylor,D., Wang,G. and Palcic,M.
JOURNAL .alpha.1,2-fucosyltransferase
FEATURES Patent: US 6670160-A 1 30-DEC-2003;
source Location/Qualifiers
1..1119
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 56.4%; Score 22; DB 6; Length 1119;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTTAATTCTA 39
|||||
Db 1098 TTATAGCGAGCGTCTTAATTCTA 1119

RESULT 12
AF093828 1122 bp DNA linear BCT 13-MAY-1999
LOCUS Helicobacter pylori strain UAI234 alpha-1,2-fucosyltransferase
DEFINITION AF093828
ACCESSION AF093828
VERSION AF093828.1 GI:4808587
KEYWORDS .
SOURCE Helicobacter pylori
ORGANISM Helicobacter pylori
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Helicobacteraceae; Helicobacter.
REFERENCE 1 (bases 1 to 1122)
AUTHORS Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.
TITLE Molecular genetic basis for the variable expression of Lewis X
antigen in Helicobacter pylori: analysis of the alpha (1,2)
fucosyltransferase gene
Mol. Microbiol. 31 (4), 1265-1274 (1999)
10096092
JOURNAL 2 (bases 1 to 1122)
REFERENCE Wang,G. and Taylor,D.E.
AUTHORS Direct Submission
TITLE Submitted (24-SEP-1998) Medical Microbiology and Immunology,
JOURNAL University of Alberta, 1-51 Medical Science Building, Edmonton, AB
T6G 2H7, Canada
FEATURES Location/Qualifiers
source 1..1122
/organism="Helicobacter pylori"
/mol_type="genomic DNA"
/strain="UAI234"
/db_xref="taxon:210"
142..1041
/codon_start=1
/transl_table=11
/product="alpha-1,2-fucosyltransferase"
/protein_id="AAD29863.1"
/db_xref="GI:4808588"
/translation="MAKVVQICGGLGNOMFOYAPAKSLQKHSNTPVLDDITSPDWSN
RKMQLEFPIDLPYASKEIALAKMQLPKLVRNVLCMGDFRVSQELVFPYEPKLLK
TSRLTYFYQDPDRYDAISPLIKQFTLPFPPEPENGNNKKKBEYHRKGLALIAKN
SVFVHRRGDYVGICQIDYOKKALEYMAKRVPMELFVFCDELEFTLDLGYPF
MDMTTRDKREAYWMLIMQSKGIIANSTYSWAAYLINNPKIIGPKHWFGEH
NILCKEWKIESHFVKSQYNA"

CDS
ORIGIN
Query Match 56.4%; Score 22; DB 1; Length 1122;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	18	TTATAGCGAGCGTCTAAATTCTA	39
Db	1101	TTATAGCGAGCGTCTAAATTCTA	1122
RESULT 13			
AF093829	LOCUS	AF093829	1124 bp DNA linear BCT 13-MAY-1999
DEFINITION	Helicobacter pylori strain UAI182 alpha-1,2-fucosyltransferase gene, alternative products, complete cds.		
ACCESSION	AF093829	VERSION	AF093829.1 GI:4808589
KEYWORDS	Helicobacter pylori		
SOURCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 1124)		
AUTHORS	Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.		
TITLE	Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori: analysis of the alpha (1,2) fucosyltransferase gene		
JOURNAL	Mol. Microbiol.	31 (4),	1265-1274 (1999)
PUBMED	10096092		
REFERENCE	2 (bases 1 to 1124)		
AUTHORS	Wang,G. and Taylor,D.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-SEP-1998) Medical Microbiology and Immunology, University of Alberta, 1-51 Medical Science Building, Edmonton, AB T6G 2H7, Canada		
FEATURES	source Location/Qualifiers		
	1..1124		
	/organism="Helicobacter pylori"		
	/mol_type="genomic DNA"		
	/strain="UAI182"		
	/db_xref="taxon:210"		
	join(141..573,573..1042)		
	/note="ribosomal slippage"		
	/codon_start=1		
	/transl_table=11		
	/product="alpha-1,2-fucosyltransferase long form"		
	/protein_id="AAD29865.1"		
	/db_xref="GI:4808591"		
	/translation="MAFKVVRICGGNGNQMFQYAPAKSLQKHSNTPVLDDITSPDWSN RMQLLEPFDLPYASAKEIAIKMHLPKLVDRALVKYIGFDRVSQIVFEYEPKLK PSRLTYFGYFQDPFRYPDAISPLIKQTFTLPPPKIIRIIKKEEYHRKLAILAAK NSVFHIRRGDYGVGCQIDGYQKALEYAKRVPMNELPFVCFEDLTFTQNLDLGYP FMDTTRDKDEARYWMLLMQSCQHGIANTSTYSWWAAAYLINNPBKIIIGPKHMLPGFH ENILCKEWKIESHFVEKSQKNA"		
	141..605		
	/note="nonfunctional"		
	/codon_start=1		
	/transl_table=11		
	/product="alpha-1,2-fucosyltransferase short form"		
	/protein_id="AAD29864.1"		
	/db_xref="GI:4808590"		
	/translation="MAFKVVRICGGNGNQMFQYAPAKSLQKHSNTPVLDDITSPDWSN RMQLLEPFDLPYASAKEIAIKMHLPKLVDRALVKYIGFDRVSQIVFEYEPKLK PSRLTYFGYFQDPFRYPDAISPLIKQTFTLPPPKIIRIIKKKNNTANLL"		
CDS			
ORIGIN			
Query Match	56.4%;	Score 22;	DB 1; Length 1124;
Best Local Similarity	100.0%;	Pred. No. 52;	
Matches	22;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	18	TTATAGCGAGCGTCTAAATTCTA	39
Db	1103	TTATAGCGAGCGTCTAAATTCTA	1124
RESULT 14			
AF093832	LOCUS	AF093832	1124 bp DNA linear BCT 13-MAY-1999
DEFINITION	Helicobacter pylori strain UAI210 alpha-1,2-fucosyltransferase gene, complete cds.		
ACCESSION	AF093832	VERSION	AF093832.1 GI:4808598
KEYWORDS	Helicobacter pylori		
SOURCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 1124)		
AUTHORS	Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.		
TITLE	Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori: analysis of the alpha (1,2) fucosyltransferase gene		
JOURNAL	Mol. Microbiol.	31 (4),	1265-1274 (1999)
PUBMED	10096092		
REFERENCE	2 (bases 1 to 1124)		
AUTHORS	Wang,G. and Taylor,D.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-SEP-1998) Medical Microbiology and Immunology, University of Alberta, 1-51 Medical Science Building, Edmonton, AB T6G 2H7, Canada		
FEATURES	source Location/Qualifiers		
	1..1124		
	/organism="Helicobacter pylori"		
	/mol_type="genomic DNA"		
	/strain="UAI182"		
	/db_xref="taxon:210"		
	join(141..573,573..1042)		
	/note="ribosomal slippage"		
	/codon_start=1		
	/transl_table=11		
	/product="alpha-1,2-fucosyltransferase long form"		
	/protein_id="AAD29865.1"		
	/db_xref="GI:4808591"		
	/translation="MAFKVVRICGGNGNQMFQYAPAKSLQKHSNTPVLDDITSPDWSN RMQLLEPFDLPYASAKEIAIKMHLPKLVDRALVKYIGFDRVSQIVFEYEPKLK PSRLTYFGYFQDPFRYPDAISPLIKQTFTLPPPKIIRIIKKEEYHRKLAILAAK NSVFHIRRGDYGVGCQIDGYQKALEYAKRVPMNELPFVCFEDLTFTQNLDLGYP FMDTTRDKDEARYWMLLMQSCQHGIANTSTYSWWAAAYLINNPBKIIIGPKHMLPGFH ENILCKEWKIESHFVEKSQKNA"		
	141..605		
	/note="nonfunctional"		
	/codon_start=1		
	/transl_table=11		
	/product="alpha-1,2-fucosyltransferase short form"		
	/protein_id="AAD29864.1"		
	/db_xref="GI:4808590"		
	/translation="MAFKVVRICGGNGNQMFQYAPAKSLQKHSNTPVLDDITSPDWSN RMQLLEPFDLPYASAKEIAIKMHLPKLVDRALVKYIGFDRVSQIVFEYEPKLK PSRLTYFGYFQDPFRYPDAISPLIKQTFTLPPPKIIRIIKKKNNTANLL"		
CDS			
ORIGIN			
Query Match	56.4%;	Score 22;	DB 1; Length 1124;
Best Local Similarity	100.0%;	Pred. No. 52;	
Matches	22;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	18	TTATAGCGAGCGTCTAAATTCTA	39
Db	1103	TTATAGCGAGCGTCTAAATTCTA	1124
RESULT 14			
AF093832	LOCUS	AF093832	1124 bp DNA linear BCT 13-MAY-1999
DEFINITION	Helicobacter pylori strain UAI210 alpha-1,2-fucosyltransferase gene, complete cds.		
ACCESSION	AF093832	VERSION	AF093832.1 GI:4808598
KEYWORDS	Helicobacter pylori		
SOURCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter.		
REFERENCE	1 (bases 1 to 1124)		
AUTHORS	Wang,G., Rasko,D.A., Sherburne,R. and Taylor,D.E.		
TITLE	Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori: analysis of the alpha (1,2) fucosyltransferase gene		
JOURNAL	Mol. Microbiol.	31 (4),	1265-1274 (1999)
PUBMED	10096092		
REFERENCE	2 (bases 1 to 1124)		
AUTHORS	Wang,G. and Taylor,D.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-SEP-1998) Medical Microbiology and Immunology, University of Alberta, 1-51 Medical Science Building, Edmonton, AB T6G 2H7, Canada		
FEATURES	source Location/Qualifiers		
	1..1124		
	/organism="Helicobacter pylori"		
	/mol_type="genomic DNA"		
	/strain="UAI207"		
	/db_xref="taxon:210		

JOURNAL Submitted (24-SEP-1998) Medical Microbiology and Immunology,  
University of Alberta, 1-51 Medical Science Building, Edmonton, AB  
T6G 2H7, Canada

FEATURES Location/Qualifiers  
source 1. .1124  
/organism="Helicobacter pylori"  
/mol\_type="genomic DNA"  
/strain="UA1210"  
/db\_xref="taxon:210"  
143. .1045  
CDS /codon\_start=1  
/transl\_table=11  
/product="alpha-1,2-fucosyltransferase"  
/protein\_id="AAD29869.1"  
/db\_xref="GI:4808599"  
/translation="MAFKVVOICGGLGNQMFQYAFAKSLQKHSNTPVLLDITSPDWSD  
RKMQLLEFPINLPYASAKBIAIAKMOHLPKLVLDALKCMGFDRVSOEIVPEYEPGLLK  
PSRLTYFGYQDPYFDIAISPLIKQTFTLPPEPNKNNKKEEYHRKLSLTLAAK  
NSVFVHRRGDYVGIGQLGIDYOKALEYMAKRVPMNELFVPCEDLEFQNLDLGYP  
FMDMTTRNKREAYMDLMOGCOGIIANSTYSWMAAYLIENPEKIIIGPKHWFGLGH  
ENILCKEWVKIESHFVKSQKYN"

ORIGIN

Query Match	56.4%	Score 22;	DB 1;	Length 1124;
Best Local Similarity	100.0%	Pred. No. 52;		
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39  
|||||  
Db 1103 TTATAGCGAGCGTCTAATTCTA 1124  
|||||

Search completed: January 30, 2006, 02:34:18  
Job time : 898 secs



GenCore version 5.1.6

Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:24:06 ; Search time 1556 Seconds  
(without alignment)  
1172.684 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4

Perfect score: 39

Sequence: 1 gaacactcacagcgtcttatagcagcgtcttaattctta 39

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_hic:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_est7:\*

9: gb\_ges1:\*

10: gb\_ges2:\*

11: gb\_ges3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23.8	61.0	1454	9	BZ554770
C 2	22.6	57.9	392	5	BY415162
C 3	22.6	57.9	500	6	CD599349
C 4	22	56.4	415	5	BY456864
C 5	22	56.4	599	2	BG606513
C 6	22	56.4	657	5	BQ805941
C 7	22	56.4	664	1	AJ793428
C 8	22	56.4	680	5	BQ743392
C 9	22	56.4	700	6	CD866027
C 10	22	56.4	711	6	CD490533
C 11	22	56.4	719	1	AJ796733
C 12	22	56.4	728	2	B8972908
C 13	22	56.4	744	1	AJ797420
C 14	22	56.4	762	1	AJ796856
C 15	22	56.4	836	8	CV761672
C 16	22	56.4	839	8	CV761673
C 17	22	56.4	1007	11	CNS05ET8
C 18	22	56.4	1116	7	CK207657
C 19	22	56.4	1124	8	DR736695
C 20	21.8	55.9	387	5	BY580486
C 21	21.8	55.9	399	5	BY506980
C 22	21.8	55.9	412	5	BY431624

C 23	21.8	55.9	449	8	M75881
C 24	21.8	55.9	865	2	BF572390
C 25	21.6	55.4	431	7	CV039590
C 26	21.6	55.4	494	5	B0290064
C 27	21.6	55.4	560	5	B0360261
C 28	21.6	55.4	576	1	AJ735914
C 29	21.6	55.4	595	1	AJ450530
C 30	21.6	55.4	623	1	AJ735924
C 31	21.6	55.4	630	9	BH764985
C 32	21.6	55.4	647	3	BM427220
C 33	21.6	55.4	685	7	CN233065
C 34	21.6	55.4	692	1	AJ452211
C 35	21.6	55.4	727	7	CK356301
C 36	21.6	55.4	761	5	B0387220
C 37	21.6	55.4	817	1	AJ450262
C 38	21.6	55.4	843	5	B0293792
C 39	21.6	55.4	863	5	B0324884
C 40	21.4	54.9	627	10	AG271419
C 41	21.4	54.9	680	9	CC418927
C 42	21.4	54.9	680	9	CC418930
C 43	21.4	54.9	920	9	CC357981
C 44	21.2	54.4	453	1	AA099175
C 45	21.2	54.4	741	10	BK157310

## ALIGNMENTS

RESULT 1 BZ554770 1454 bp DNA linear GSS 17-DEC-2002  
LOCUS pacsl-60\_4851.xl pacsl-60 Pseudomonas aeruginosa genomic clone  
DEFINITION pacsl-60\_4851, genomic survey sequence.  
ACCESSION BZ554770  
VERSION BZ554770.1 GI:27162705  
KEYWORDS GSS.  
SOURCE Pseudomonas aeruginosa  
ORGANISM Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.  
REFERENCES 1 (bases 1 to 1454)  
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
Burns, J.L., Kaul, R. and Olsen, M.V.  
TITLE Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
JOURNAL  
COMMENT Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
FEATURES  
source  
1. .1454  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="1-60"  
/db\_xref="taxon:287"  
/clone="pacsl-60\_4851"  
/clone\_lib="pacsl-60"  
/note="clinical isolate 1-60 Whole genomic shotgun  
library."

ORIGIN  
Query Match 61.0%; Score 23.8; DB 9; Length 1454;  
Best Local Similarity 80.0%; Pred. No. 11;  
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 4 CACTCACACGGCTTATAGCAGCGTCTATTCT 38  
| | | | | | | | | | | | | | | | | | | | | |  
Db 906 CCCTCACTTGGTCGTATCTAGCGTCTACTTCT 940

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

**FEATURES**  
**source**

```
1. 392
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I920009B12"
/tissue_type="kidney"
/dev_stage="17 days embryo"
/clone_lib="RIXEN full-length
kidney"
```

## ORIGIN

Query Match	57.9%	Score 22.6	DB 5	Length 392
Best Local Similarity	86.2%	Fed. No. 32		
Matches 25	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Y	8	CA CAG CGCTCTTATAG CGAGCGTCTAA TT	36	
b	162	CA CAG CGGACTTTATAG CGAGAGTATA TT	134	

RESULT 3	CD599349	500 bp	mRNA	linear	EST 16-JUN-2003
LOCUS	CD599349				
DEFINITION	RK115A4H02.T3 Zebrafish Kidney Marrow cDNA library Danio rerio cDNA clone RK115A4H02 5', mRNA sequence.				
ACCESSION	CD599349				
VERSION	CD599349.1	GI:31780225			
KEYWORDS	EST.				
SOURCE	Danio rerio				
ORGANISM	Danio rerio (zebrafish)				

**ORGANISM**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 500)

**REFERENCE**  
Song, H.D., Sun, X.J., Deng, M., Zhang, G.W., Zhou, Y., Wu, X.Y.,  
Sheng, Y., Chen, Y., Ruan, Z., Jiang, C.L., Fan, H.Y., Zou, L.I.,  
Kanki, J.P., Liu, T.X., Look, A.T. and Chen, Z.  
Hematopoietic gene expression profile in zebrafish kidney marrow  
Proc. Natl. Acad. Sci. U.S.A. 101 (46), 16240-16245 (2004)  
15520368  
TURNED

**FEATURES**

```
1.500
/organism="Danio rerio"
/mol_type="mrna"
/db_xref="taxon:7955"
/clone="RK115A4H02"
```

```

/clone="lyb-zebrafish Kidney Marrow cDNA library"
/notes="Organ: kidney; Vector: pBS-CMW; Site 1: XhoI;
Site 2: EcoRI; Total RNA was extracted from the kidney
tissues of mature zebrafish. The poly (A)+ RNA fraction
was separated from total RNA by oligo (dT) cellulose
chromatography. Library was initially constructed in the
lambdaZAP Express vector (Stratagene) and in vivo excised
into pBS-CMW vector."

```

## ORIGIN

BY415162 392 bp mRNA linear EST 13-DEC-2002  
 BY415162 RIKEN full-length enriched, 17 days embryo Mus  
 musculus cDNA clone 1920009812 3', mRNA sequence.  
 BY415162  
 BY415162.1 GI:26680286  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Muridae; Murinae; Mus.  
 1 (bases 1 to 392)

REFERENCE AUTHORS	TITLE
1. (bases 1 to 392)	
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Kikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasaki, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, M., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshew-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, K., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, N., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
	Nature 420, 563-573 (2002)
	12456851
	PUBLISHED

TITLE	Analysis of the mouse transcriptome based on functional annotation
JOURNAL	of 60,770 full-length cDNAs
PUBMED	Nature 420, 563-573 (2002)
COMMENT	Contact: Yoshihide Hayashizaki 12466851

Contact: toshimune mayanazaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Query Match 57.9%; Score 22.6; DB 6; Length 500;  
 Best Local Similarity 75.7%; Pred. No. 33;  
 Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAACACTCACACGGCTCTTATAGCAGCGCTCTAAATTC 37  
 |||||  
 Db 165 GAACACTCCAGGTTCTCTACAGCGAAGGCTCACTTC 201

RESULT 4  
 BY456864/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BY456864 415 bp mRNA EST 13-DEC-2002  
 BY456864 RIKEN full-length enriched, 17 days embryo whole body Mus  
 musculus cDNA clone K630094C08 3', mRNA sequence.  
 BY456864  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus

REFERENCE  
 AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,  
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,  
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,  
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,  
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,  
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,  
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,  
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W.J., Perteau, G., Pesole, G.,  
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,  
 Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,  
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,  
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, X., Watanabe, Y.,  
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,  
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,  
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,  
 Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

TITLE  
 JOURNAL  
 PUBMED  
 COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/  
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,  
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,  
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,  
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,  
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and  
 Hayashizaki, Y. Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and abstraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

FEATURES  
 source

Location/Qualifiers  
 1..415  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="K630094C08"  
 /tissue\_type="whole body"  
 /dev\_stage="17 days embryo"  
 /clone\_lib="RIKEN full-length enriched, 17 days embryo  
 whole body"

## ORIGIN

Query Match 56.4%; Score 22; DB 5; Length 415;  
 Best Local Similarity 83.3%; Pred. No. 61;  
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 8 CACACGGCTCTTATAGCAGCGCTCTAAATTC 37  
 |||||  
 Db 185 CACACGGATTATATAGCGAGAGTATAATTC 156

## RESULT 5

BG606513/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

BG606513 599 bp mRNA linear EST 17-APR-2001  
 WHE2957 B01\_D012S Wheat dormant embryo cDNA library Triticum  
 aestivum cDNA clone WHE2957\_B01\_D01, mRNA sequence.  
 BG606513  
 BG606513.1 GI:13656484  
 Triticum aestivum (bread wheat)  
 Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poideae; Triticeae; Triticum.  
 1 (bases 1 to 599)  
 Anderson, O.D., Chao, S., Chin, A., Close, T.J., Doherty, L.,  
 Fenton, R.D., Lazo, G.R., Rausch, C.J., Walker-Simmons, M.K. and  
 Wilson, C.  
 The structure and function of the expressed portion of the wheat  
 genomes - Dormant embryo cDNA library  
 Unpublished (2001)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Stratagene SK primer.

FEATURES  
 source

Location/Qualifiers  
 1..599  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Brevor"  
 /db\_xref="taxon:4565"  
 /clone="WHE2957\_B01\_D01"

TITLE  
 JOURNAL  
 PUBMED  
 COMMENT

Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanderson@pw.usda.gov  
 Sequence have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: Stratagene SK primer.

FEATURES  
 source

Location/Qualifiers  
 1..599  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Brevor"  
 /db\_xref="taxon:4565"  
 /clone="WHE2957\_B01\_D01"



```

Qy 1 GAACACTCACACGCGTCTTATAGCGAGCGTCTAAATTTCT 38
Db 54 GAGCAACCAACACGCGTGGCGGTATCGAGCGTGGAATTTCT 91

RESULT 8
BQ743392/c
LOCUS BQ743392 680 bp mRNA linear EST 17-JUL-2002
DEFINITION WHE4103 D12 H23S Wheat salt-stressed root cDNA library Triticum aestivum cDNA clone WHE4103_D12_H23, mRNA sequence.
ACCESSION BQ743392
VERSION BQ743392.1 GI:21890179
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1 (bases 1 to 680)
AUTHORS Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K., Dvorak, J., Izzo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.
TITLE The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderan@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
source
Location/Qualifiers
1..680
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4103_D12_H23"
/tissue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/clone_lib="Wheat salt-stressed root cDNA library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hygroponic plants grown to full tillering stage were treated with 150 mM NaCl for either 12 hours or 7 days. Root tissues of the plants subjected to both types of treatment were collected separately at University of California, Davis (E. Akhunov and K. Deal in J. Dvorak's Lab). Total RNA was prepared separately from the two samples (12h and 7day treatments), and equal amount of RNA was then pooled. PolyA RNA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak) at the University of California, Davis. Colony plating, plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
```

Query Match 56.4%; Score 22; DB 5; Length 680;  
Best Local Similarity 73.7%; Pred. No. 64;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Qy 2 AACACTCACACGCGTCTTATAGCGAGCGTCTAAATTTCTA 39
Db 172 AACACTCACGCGTCTTATAGCGAGCGTCTTAAATTTCTA 135

RESULT 9
BQ743392/c
LOCUS BQ743392 680 bp mRNA linear EST 17-JUL-2002
DEFINITION WHE4103 D12 H23S Wheat salt-stressed root cDNA library Triticum aestivum cDNA clone WHE4103_D12_H23, mRNA sequence.
ACCESSION BQ743392
VERSION BQ743392.1 GI:21890179
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1 (bases 1 to 680)
AUTHORS Anderson, O.D., Akhunov, E., Chao, S., Crossman, C., Deal, K., Dvorak, J., Izzo, G.R., Pham, J., Rausch, C.J., Wilson, C. and Woo, J.
TITLE The structure and function of the expressed portion of the wheat genomes - Salt-stressed root cDNA library
JOURNAL Unpublished (2002)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderan@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: SK primer.

FEATURES
source
Location/Qualifiers
1..680
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4103_D12_H23"
/tissue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/clone_lib="Wheat salt-stressed root cDNA library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Hygroponic plants grown to full tillering stage were treated with 150 mM NaCl for either 12 hours or 7 days. Root tissues of the plants subjected to both types of treatment were collected separately at University of California, Davis (E. Akhunov and K. Deal in J. Dvorak's Lab). Total RNA was prepared separately from the two samples (12h and 7day treatments), and equal amount of RNA was then pooled. PolyA RNA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript SK(-) phagemids in J. Dvorak's lab (E. Akhunov, J. Dvorak) at the University of California, Davis. Colony plating, plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
```

Query Match 56.4%; Score 22; DB 5; Length 680;  
Best Local Similarity 73.7%; Pred. No. 64;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

CD866027/c
LOCUS CD866027 700 bp mRNA linear EST 11-JUL-2003
DEFINITION AZO2.102G05F001123 AZO2 Triticum aestivum cDNA clone AZO2102G05, mRNA sequence.
ACCESSION CD866027
VERSION CD866027.1 GI:32549843
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1 (bases 1 to 700)
AUTHORS Genoplante, a major partnership french program in plant genomics
TITLE Unpublished (2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).

FEATURES
source
Location/Qualifiers
1..700
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO2102G05"
/tissue_type="root"
/clone_lib="AZO2"

ORIGIN
Query Match 56.4%; Score 22; DB 6; Length 700;  

Best Local Similarity 73.7%; Pred. No. 64;  

Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```

Qy 2 AACACTCACACGCGTCTTATAGCGAGCGTCTAAATTTCTA 39
Db 163 AACACTCACGCGTCTTATAGCGAGCGTCTTAAATTTCTA 126

RESULT 10
CD490533
LOCUS CD490533 711 bp mRNA linear EST 05-JUN-2003
DEFINITION WHE2957_B01_D012T Brevor wheat dormant embryo cDNA library Triticum aestivum cDNA clone WHE2957_B01_D01, mRNA sequence.
ACCESSION CD490533
VERSION CD490533.1 GI:31417579
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1 (bases 1 to 711)
AUTHORS Anderson, O.D., Chao, S., Chin, A., Close, T.J., Crossman, C., Doherty, L., Fenton, R.D., Izzo, G.R., Rausch, C.J., Walker-Simmons, M.K. and Wilson, C.
TITLE The structure and function of the expressed portion of the wheat genomes - Brevor dormant embryo cDNA library
JOURNAL Unpublished (2003)
COMMENT Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderan@pw.usda.gov
This EST was generated by sequencing from the 3' end of the clone. Sequences have been trimmed to remove vector sequence and low
```

quality sequence with phred score less than 20.  
Seq primer: T7 primer.

# FEATURES

Location/Qualifiers  
1..711

/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="Brevor"  
/db\_xref="taxon:4565"  
/clone="WHE2957 B01.D01"  
/tissue\_type="Seed embryo"  
/dev\_stage="Mature seed"  
/lab\_host="E. coli SOLR"  
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown to seed maturity under conditions favoring seed dormancy (L. Doherty at K. Walker-Simmons lab, Washington State University, Pullman, WA). Embryos were cut from mature dormant seed (Doherty). Total RNA was prepared from these embryos, polyA was purified, a cDNA library was made, and the cDNA clones were in vivo excised to give pBluescript phagemids in the TJ clone lab at the University of California, Riverside (Chin, Fenton). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

## ORIGIN

Query Match 56.4%; Score 22; DB 6; Length 711;

Best Local Similarity 73.7%; Pred. No. 64;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 AACACTCACACGGCTCTTATAGCGAGCGTCTAATTCTA 39

Db 608 AACACTCACACGGCTCTTATAGCGAGCGTCTAATTCTA 645

## RESULT 11

LOCUS AJ796733 719 bp mRNA linear EST 08-DEC-2004  
DEFINITION AJ796733 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018.3.11.n14, mRNA sequence.

ACCESSION AJ796733

VERSION AJ796733.1 GI:51112061

KEYWORDS EST.

SOURCE Antirrhinum majus (snapdragon)

ORGANISM Antirrhinum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Plantaginales; Antirrhineae; Antirrhinum.

REFERENCE 1 (bases 1 to 719)

AUTHORS Bey, M., Stueber, K., Fellenberg, K., Schwarz-Sommer, Z., Sommer, H., Saedler, H., and Zachgo, S.

TITLE Characterization of Antirrhinum Petal Development and

Identification of Target Genes of the Class B MADS Box Gene

DEFICIENS

JOURNAL Plant Cell 16 (12), 3197-3215 (2004)

PUBMED 15539471

COMMENT Contact: Schwarz-Sommer Z

MPI fuer Zuechtungsforschung

Carl-von-Linne Weg 10, D-50829, Germany.

Location/Qualifiers

1..719

/organism="Antirrhinum majus"

/mol\_type="mRNA"

/db\_xref="taxon:4151"

/clone="018.3.11.n14"

/tissue\_type="whole plant"

/clone\_lib="Antirrhinum majus whole plant"

## ORIGIN

Query Match

56.4%; Score 22; DB 1; Length 719;

Best Local Similarity 73.7%; Pred. No. 64;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GAACACTCACACGGCTCTTATAGCGAGCGTCTAATTCT 38

Db 61 GAGCAACACACACGGCTCGCGTATCGAGCGTGAATTTCT 98

## RESULT 12

BE972908/c

LOCUS BE972908

DEFINITION BE972908 728 bp mRNA linear EST 04-OCT-2000  
mRNA sequence.

ACCESSION BE972908

VERSION BE972908.1 GI:10586244

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 728)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILLNL at:

<http://image.llnl.gov>

Plate: LICM778 row: h column: 13

High quality sequence stop: 2.

## FEATURES

source

1..728

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3935676"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_lib="NIH MGC 82"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:

SfiI (ggcccttcggcc); Site 2: SfiI (ggccattatggcc); 5' and

3' adaptors were used in cloning as follows: 5' adaptor

sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:

5'-ATTCTAGAGCGCGCGGCACATG-dt(30)BN-3' (where B = A,

C, or G and N = A, C, or T). Average insert size

1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained

inserts by PCR. This library was enriched for full-length

clones and was constructed by Clontech Laboratories (Palo

Alto, CA)."

## ORIGIN

Query Match 56.4%; Score 22; DB 2; Length 728;

Best Local Similarity 73.7%; Pred. No. 64;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GAACACTCACACGGCTCTTATAGCGAGCGTCTAATTCT 38

Db 357 GAACACTCATCTGCTCTTTAAGAGAGAGTGTGATTAT 320

## RESULT 13

AJ797420

LOCUS AJ797420

DEFINITION AJ797420 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018.4.01.m23, mRNA sequence.

ACCESSION AJ797420

VERSION AJ797420.1 GI:51112748

KEYWORDS EST.

```

SOURCE      Antirrhinum majus (snapdragon)
ORGANISM    Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1 (bases 1 to 744)
AUTHORS    Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H.,
            Saedler,H. and Zachgo,S.
TITLE      Characterization of Antirrhinum Petal Development and
            Identification of Target Genes of the Class B MADS Box Gene
            DFIC1ENS
JOURNAL    Plant Cell 16 (12), 3197-3215 (2004)
PUBMED     15539471
COMMENT    Contact: Schwarz-Sommer Z
            Molekulare Pflanzen-genetik
            MPI fuer Zuechtungs-forschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES   source
            1..744
            /organism="Antirrhinum majus"
            /mol_type="mRNA"
            /db_xref="taxon:4151"
            /clone="018_4_01_m23"
            /tissue_type="whole plant"
            /clone_lib="Antirrhinum majus whole plant"

ORIGIN
Query Match      56.4%; Score 22; DB 1; Length 744;
Best Local Similarity 73.7%; Pred. No. 65;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy  1 GAACACTCACACGGCTCTTATAGCAGCGCTCAATTCT 38
    ||| ||||| ||| ||| ||| ||| ||| ||| |||
Db  62 GAGCAACACACGGCTGCCGTATCGAGCGTCAAAATTCT 99

RESULT 14
AJ796856
LOCUS      Antirrhinum majus whole plant Antirrhinum majus cDNA clone
DEFINITION
AJ796856 762 bp mRNA linear EST 08-DEC-2004
O18_3_12_c22, mRNA sequence.
ACCESSION  AJ796856
VERSION     AJ796856.1 GI:51112184
KEYWORDS    EST.
SOURCE      Antirrhinum majus (snapdragon)
ORGANISM    Antirrhinum majus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            asterids; lamids; Lamiales; Plantaginaceae; Antirrhineae;
            Antirrhinum.
REFERENCE   1 (bases 1 to 762)
AUTHORS    Bey,M., Stueber,K., Fellenberg,K., Schwarz-Sommer,Z., Sommer,H.,
            Saedler,H. and Zachgo,S.
TITLE      Characterization of Antirrhinum Petal Development and
            Identification of Target Genes of the Class B MADS Box Gene
            DFIC1ENS
JOURNAL    Plant Cell 16 (12), 3197-3215 (2004)
PUBMED     15539471
COMMENT    Contact: Schwarz-Sommer Z
            Molekulare Pflanzen-genetik
            MPI fuer Zuechtungs-forschung
            Carl-von-Linne Weg 10, D-50829, Germany.
FEATURES   source
            1..762
            /organism="Antirrhinum majus"
            /mol_type="mRNA"
            /db_xref="taxon:4151"
            /clone="018_3_12_c22"
            /tissue_type="whole plant"
            /clone_lib="Antirrhinum majus whole plant"

ORIGIN

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:49:51 ; Search time 73 Seconds  
(without alignments)  
949.656 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4

Perfect score: 39  
Sequence: 1 gaacactacacgcgtcttatagcagcgcttaattcta 39

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgm2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgm2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgm2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgm2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgm2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgm2\_6/ptodata/1/ina/PCUTS COMB.seq.\*  
7: /cgm2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgm2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgm2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	56.4	1119	3	US-09-433-598-1
2	22	56.4	1119	3	US-09-848-838A-1
3	20	51.3	20	3	US-09-433-598-4
4	20	51.3	20	3	US-09-848-838A-4
5	19.2	49.2	2664	3	US-09-489-039A-4547
6	19.2	49.2	36899	3	US-09-949-016-13566
7	19	48.7	19	3	US-09-433-598-3
8	19	48.7	19	3	US-09-848-838A-3
9	19	48.7	2670	3	US-09-489-039A-70
10	19	48.7	124884	3	US-09-661-596A-76
11	19	48.7	124884	3	US-09-913-514-1
12	19	48.7	124884	3	US-10-288-823-76
13	19	48.7	125157	3	US-09-913-514-2
14	18.8	48.2	537	3	US-09-710-279-1175
15	18.8	48.2	563	3	US-09-495-050A-210
16	18.8	48.2	601	3	US-09-949-002-3268
17	18.8	48.2	601	3	US-09-949-002-3269
18	18.8	48.2	601	3	US-09-949-002-10786
19	18.8	48.2	601	3	US-09-949-002-10787
20	18.8	48.2	730	3	US-09-533-559-868
21	18.8	48.2	2226	3	US-09-710-279-1177
22	18.8	48.2	2301	3	US-09-134-001C-852
23	18.8	48.2	3511	3	US-09-710-279-3461
24	18.8	48.2	3705	3	US-09-710-279-3592

25	18.8	48.2	34725	3	US-09-949-002-665	Sequence 665, App
26	18.8	48.2	34726	3	US-09-949-002-857	Sequence 857, App
c 27	18.8	48.2	35688	3	US-09-949-016-16873	Sequence 16873, A
28	18.8	48.2	49795	3	US-09-453-702B-60	Sequence 60, Appl
29	18.8	48.2	49795	3	US-10-114-170-60	Sequence 60, Appl
30	18.8	48.2	236341	3	US-09-949-016-13978	Sequence 13978, A
31	18.8	48.2	786431	3	US-09-751-389-3	Sequence 3, Appl
32	18.6	47.7	601	3	US-09-949-016-135264	Sequence 135264, A
c 33	18.6	47.7	107421	3	US-09-949-016-15532	Sequence 15532, A
c 34	18.4	47.2	601	3	US-09-949-016-168281	Sequence 168281, A
35	18.4	47.2	1397	3	US-09-270-767-5652	Sequence 5652, Ap
36	18.4	47.2	1397	3	US-09-270-767-20934	Sequence 20934, A
c 37	18.4	47.2	2394	3	US-09-861-451A-11	Sequence 11, Appl
c 38	18.4	47.2	2661	3	US-10-104-047-38	Sequence 38, Appl
c 39	18.4	47.2	60589	3	US-09-949-016-17070	Sequence 17070, A
c 40	18.4	47.2	116652	3	US-09-949-016-13413	Sequence 13413, A
c 41	18.4	47.2	225127	3	US-09-949-016-16480	Sequence 16480, A
42	18.2	46.7	601	3	US-09-949-016-24751	Sequence 24751, A
c 43	18.2	46.7	601	3	US-09-949-016-35121	Sequence 35121, A
c 44	18.2	46.7	1188	3	US-09-090-044B-1	Sequence 1, Appl
c 45	18.2	46.7	1323	3	US-09-533-559-5815	Sequence 5815, Ap

#### ALIGNMENTS

RESULT 1  
US-09-433-598-1  
; Sequence 1, Application US/09433598  
; Patent No. 6238894  
; GENERAL INFORMATION:  
; APPLICANT: Taylor Dr., Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: Alphas, 2 Fucosyltransferase  
; FILE REFERENCE: 07254/061001  
; CURRENT APPLICATION NUMBER: US/09/433,598  
; EARLIER FILING DATE: 1998-11-02  
; EARLIER FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (137)..(1036)  
US-09-433-598-1

Query Match 56.4%; Score 22; DB 3; Length 1119;  
Best Local Similarity 100.0%; Pred.No.2.4; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 18 TTATAGCGAGCGTCTTAATTCTA 39  
|||||  
Db 1098 TTATAGCGAGCGTCTTAATTCTA 1119

RESULT 2  
US-09-848-838A-1  
; Sequence 1, Application US/09848838A  
; Patent No. 6670160  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
; FILE REFERENCE: 07254-061002  
; CURRENT APPLICATION NUMBER: US/09/848,838A  
; EARLIER FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 09/433,598

; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: US 60/107,268  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (137)...(1036)  
US-09-848-838A-1

Query Match 56.4%; Score 22; DB 3; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 2.4;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGGCGTCTAAATTTCTA 39  
|||||  
Db 1098 TTATAGCGGCGTCTAAATTTCTA 1119

## RESULT 3

US-09-433-598-4/c  
; Sequence 4, Application US/09433598  
; Patent No. 6238894  
; GENERAL INFORMATION:  
; APPLICANT: Taylor Dr., Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase  
; FILE REFERENCE: 07254/061001  
; CURRENT APPLICATION NUMBER: US/09/433,598  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: 60/107268  
; EARLIER FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: oligonucleotide  
US-09-433-598-4

Query Match 51.3%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATAGCGAGCGTCTAAATTTCTA 39  
|||||  
Db 20 ATAGCGAGCGTCTAAATTTCTA 1

## RESULT 4

US-09-848-838A-4/c  
; Sequence 4, Application US/09848838A  
; Patent No. 6670160  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
; FILE REFERENCE: 07254-061002  
; CURRENT APPLICATION NUMBER: US/09/848,838A  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 09/433,598  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: US 60/107,268  
; PRIOR FILING DATE: 1998-11-04

; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-09-848-838A-4

Query Match 51.3%; Score 20; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATAGCGAGCGTCTAAATTTCTA 39  
|||||  
Db 20 ATAGCGAGCGTCTAAATTTCTA 1

## RESULT 5

US-09-489-039A-4547/c  
; Sequence 4547, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4547  
; LENGTH: 2664  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4547

Query Match 49.2%; Score 19.2; DB 3; Length 2664;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 8 CACAGCGTCTTATAGCGAGCGTCTAAATTTCTA 39  
|||||  
Db 375 CAGCGCATCTGATAGCGCGCGGCATTCGA 344

## RESULT 6

US-09-949-016-13566  
; Sequence 13566, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13566  
; LENGTH: 36899  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature

; LOCATION: (1)....(36899)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13566

Query Match 49.2%; Score 19.2; DB 3; Length 36899;  
Best Local Similarity 75.0%; Pred. No. 90;  
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 ACACCTCACACGGCTCTTATAGCGAGGCTCTAA 34  
|||||  
DB 11789 ACACCTCACACATAACTTATAGTAGCAATTAA 11820  
|||||

## RESULT 7

US-09-433-598-3  
; Sequence 3, Application US/09433598  
; Patent No. 6238894  
; GENERAL INFORMATION:  
; APPLICANT: Taylor Dr., Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase  
; FILE REFERENCE: 07254/061001  
; CURRENT APPLICATION NUMBER: US/09/433,598  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER FILING DATE: 1998-11-02  
; EARLIER FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: oligonucleotide  
US-09-433-598-3

Query Match 48.7%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACACTCACACGGCTCTT 19  
|||||  
DB 1 GAACACTCACACGGCTCTT 19  
|||||

## RESULT 8

US-09-848-838A-3  
; Sequence 3, Application US/09848838A  
; Patent No. 6670160  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
; FILE REFERENCE: 07254-061002  
; CURRENT APPLICATION NUMBER: US/09/848,838A  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 09/433,598  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: US 60/107,268  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-09-848-838A-3

Query Match 48.7%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACACTCACACGGCTCTT 19  
|||||  
DB 1 GAACACTCACACGGCTCTT 19  
|||||

## RESULT 9

US-09-489-039A-70  
; Sequence 70, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 70  
; LENGTH: 2670  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-70

Query Match 48.7%; Score 19; DB 3; Length 2670;  
Best Local Similarity 71.4%; Pred. No. 68;  
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 ACACCTCACACGGCTCTTATAGCGAGGCTCTAATTC 37  
|||||  
DB 1295 ACACCTGGCAGATCTTAAATGCGTAGTCGAATTC 1329  
|||||

## RESULT 10

US-09-661-596A-76/c  
; Sequence 76, Application US/09661596A  
; Patent No. 6528066  
; GENERAL INFORMATION:  
; APPLICANT: Grose, Charles  
; APPLICANT: Santos, Richard  
; TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE  
; FILE REFERENCE: 140.0011 0101  
; CURRENT APPLICATION NUMBER: US/09/661,596A  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US 60/153,779  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 76  
; LENGTH: 124884  
; TYPE: DNA  
; ORGANISM: Varicella zoster  
US-09-661-596A-76

Query Match 48.7%; Score 19; DB 3; Length 124884;  
Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 AACCTCACACGGCTCTTATAGCGAGGCTCTAATT 36  
|||||  
DB 2483 AAACACACACACGACGCTGACCGAAGCTTAAATT 2449  
|||||

## RESULT 11

US-09-913-514-1/c  
; Sequence 1, Application US/09913514  
; Patent No. 6653069  
; GENERAL INFORMATION:



```

; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 210
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2378406CT1
US-09-495-050A-210

```

```

Query Match      48.2%; Score 18.8; DB 3; Length 563;
Best Local Similarity 68.4%; Pred. No. 62;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

```

QY      2 AACACTCACACGGCTTTATAGCGAGCGTCTAATTCTA 39
      |||||
Db      313 AAAGTACACATTAGTCTTCTAGCTAATGACTAATGCTA 350

```

```

Search completed: January 30, 2006, 03:28:46
Job time : 74 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 02:34:27 ; Search time 349 Seconds  
(without alignments)  
924.085 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4  
Perfect score: 39  
Sequence: 1 gaacactcacgcgtcttatagcagcgcttaattcta 39

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 973542 seqs, 4134689005 residues  
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	56.4	934	3	US-09-895-913A-253
2	22	56.4	1119	3	US-09-848-838-1
3	22	56.4	1119	7	US-10-663-033-1
4	21.8	55.9	2295	5	US-10-128-714-2164
5	21.8	55.9	2450	5	US-10-128-714-1164
6	21.8	55.9	2505	5	US-10-128-714-6164
7	21.8	55.9	2505	5	US-10-128-714-7164
8	21.8	55.9	2450	5	US-10-128-714-164
9	21.8	55.9	4505	5	US-10-128-714-5164
C 10	20.6	52.8	852	5	US-10-027-632-120434
C 11	20.6	52.8	852	5	US-10-027-632-120435
C 12	20.6	52.8	852	6	US-10-027-632-120434
C 13	20.6	52.8	852	6	US-10-027-632-120435
C 14	20.2	51.8	262	8	US-10-674-124A-12533
C 15	20.2	51.8	331	8	US-10-674-124A-12532
C 16	20.2	51.8	2454	7	US-10-282-122A-36364
C 17	20.2	51.8	29773	3	US-09-984-425-262
C 18	20.2	51.8	29773	6	US-10-158-034-99
C 19	20	51.3	20	7	US-09-848-838-4
C 20	20	51.3	20	7	US-10-663-033-4
C 21	20	51.3	117754	5	US-10-087-192-28
C 22	19.8	50.8	519	4	US-09-925-065A-415522
C 23	19.8	50.8	2349	9	US-10-795-159-694

24	19.8	50.8	3469	9	US-10-795-159-449	Sequence 449, App
25	19.8	50.8	191996	9	US-10-795-159-683	Sequence 683, App
C 26	19.8	50.8	263852	8	US-10-812-232-6	Sequence 6, Appli
C 27	19.6	50.3	600	9	US-10-972-079-66530	Sequence 66530, A
C 28	19.6	50.3	600	9	US-10-972-079-66531	Sequence 66531, A
C 29	19.6	50.3	189158	5	US-10-087-192-415	Sequence 415, App
C 30	19.6	50.3	192673	8	US-10-331-053-1	Sequence 1, Appli
C 31	19.4	49.7	546	5	US-10-027-632-215033	Sequence 215033,
C 32	19.4	49.7	546	6	US-10-027-632-215033	Sequence 215033,
C 33	19.4	49.7	560	4	US-09-925-065A-438373	Sequence 438373,
C 34	19.4	49.7	560	4	US-09-925-065A-438374	Sequence 438374,
C 35	19.4	49.7	576	7	US-10-332-859-230	Sequence 230, App
C 36	19.4	49.7	599	4	US-09-925-065A-584840	Sequence 584840,
C 37	19.4	49.7	618	5	US-10-027-632-179694	Sequence 179694,
C 38	19.4	49.7	627	5	US-10-027-632-291443	Sequence 291443,
C 39	19.4	49.7	627	5	US-10-027-632-291443	Sequence 291443,
C 40	19.4	49.7	627	5	US-10-027-632-291443	Sequence 291443,
C 41	19.4	49.7	653	3	US-09-988-067B-69	Sequence 69, Appl
C 42	19.4	49.7	685	3	US-09-930-213-103	Sequence 103, App
C 43	19.4	49.7	1695	4	US-09-925-065A-719835	Sequence 719835,
C 44	19.4	49.7	2450	6	US-10-388-934-599	Sequence 599, App
C 45	19.4	49.7	5758	3	US-09-866-050A-605	Sequence 605, App

ALIGNMENTS

RESULT 1  
US-09-895-913A-253/c  
; Sequence 253, Application US/09895913A  
; GENERAL INFORMATION:  
; APPLICANT: Kleanthous, Harold  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Miller, Charles  
; APPLICANT: Tomb, Jean Francois  
; APPLICANT: Omen, Raymond P.  
; TITLE OF INVENTION: Identification of Polynucleotides  
; TITLE OF INVENTION: Encoding No. US20020160456A1el Helicobacter Polypeptides in the  
; FILE REFERENCE: 06132/043002  
; CURRENT APPLICATION NUMBER: US/09/895,913A  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 08/881,227  
; PRIOR FILING DATE: 1997-06-24  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 253  
; LENGTH: 934  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (51)...(881)  
US-09-895-913A-253  
Query Match 56.4%; Score 22; DB 3; Length 934;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 18 TTATAGCAGCGTCTTAATTCTA 39  
|||||  
DB 844 TTATAGCAGCGTCTTAATTCTA 823  
RESULT 2  
US-09-848-838-1  
; Sequence 1, Application US/09848838  
; Patent No. US20020037570A1  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica

```
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838
; CURRENT FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)...(1036)
US-09-848-838-1
```

```
Query Match          56.4%; Score 22; DB 3; Length 1119;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 18 TTATAGCGAGCGTCTTAATCTA 39
      |||||
Db 1098 TTATAGCGAGCGTCTTAATCTA 1119
```

```
RESULT 3
US-10-663-033-1
; Sequence 1, Application US/10663033
; Publication No. US20040048331A1
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Paicic, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/10/663,033
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: US/09/848,838A
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)...(1036)
US-10-663-033-1
```

```
Query Match          56.4%; Score 22; DB 7; Length 1119;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 18 TTATAGCGAGCGTCTTAATCTA 39
      |||||
Db 1098 TTATAGCGAGCGTCTTAATCTA 1119
```

```
RESULT 4
US-10-128-714-2164
; Sequence 2164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
```

```
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2164
; LENGTH: 2295
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-2164
```

```
Query Match          55.9%; Score 21.8; DB 5; Length 2295;
Best Local Similarity 78.8%; Pred. No. 19;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
Qy 3 ACACCTCACACGCGTCTTATAGCGAGCGTCTAAT 35
      |||||
Db 1546 ACACGCAACGCGTCTTATAGCGAGCGAGATAT 1578
```

```
RESULT 5
US-10-128-714-1164
; Sequence 1164, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1164
; LENGTH: 2450
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-1164
```

```
Query Match          55.9%; Score 21.8; DB 5; Length 2450;
Best Local Similarity 78.8%; Pred. No. 19;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```



QY 3 ACACCTCACGCGTCTTATAGCGGCTCTAAT 35  
|||||  
Db 1546 ACACGCAACGCGTCTTATAGCGGAGGATAAT 1578

## RESULT 6

US-10-128-714-6164  
; Sequence 6164, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6164  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: *Aspergillus fumigatus*  
US-10-128-714-6164

Query Match 55.9%; Score 21.8; DB 5; Length 2505;  
Best Local Similarity 78.8%; Pred. No. 19;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 ACACCTCACGCGTCTTATAGCGGCTCTAAT 35  
|||||  
Db 1471 ACACGCAACGCGTCTTATAGCGGAGGATAAT 1503

## RESULT 7

US-10-128-714-7164  
; Sequence 7164, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7164  
; LENGTH: 2505  
; TYPE: DNA  
; ORGANISM: *Aspergillus fumigatus*  
US-10-128-714-7164

Query Match 55.9%; Score 21.8; DB 5; Length 2505;  
Best Local Similarity 78.8%; Pred. No. 19;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 ACACCTCACGCGTCTTATAGCGGCTCTAAT 35  
|||||  
Db 1471 ACACGCAACGCGTCTTATAGCGGAGGATAAT 1503

## RESULT 8

US-10-128-714-164  
; Sequence 164, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10/128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 164  
; LENGTH: 4450  
; TYPE: DNA  
; ORGANISM: *Aspergillus fumigatus*  
US-10-128-714-164

Query Match 55.9%; Score 21.8; DB 5; Length 4450;  
Best Local Similarity 78.8%; Pred. No. 21;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 ACACCTCACGCGTCTTATAGCGGCTCTAAT 35  
|||||  
Db 2546 ACACGCAACGCGTCTTATAGCGGAGGATAAT 2578

## RESULT 9

US-10-128-714-5164  
; Sequence 5164, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Eroshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M

;; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
;; FILE REFERENCE: 10182-018-999  
;; CURRENT APPLICATION NUMBER: US/10/128,714  
;; CURRENT FILING DATE: 2002-04-23  
;; PRIOR FILING DATE: 2001-04-23  
;; PRIOR APPLICATION NUMBER: US 60/285,697  
;; PRIOR FILING DATE: 2001-04-23  
;; PRIOR APPLICATION NUMBER: US 60/287,066  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: US 60/295,890  
;; PRIOR FILING DATE: 2001-06-05  
;; PRIOR APPLICATION NUMBER: US 60/303,899  
;; PRIOR FILING DATE: 2001-07-09  
;; PRIOR APPLICATION NUMBER: US 60/316,362  
;; PRIOR FILING DATE: 2001-08-31  
;; NUMBER OF SEQ ID NOS: 8603  
;; SOFTWARE: Patent in version 3.1  
;; SEQ ID NO 5164  
;; LENGTH: 4505  
;; TYPE: DNA  
;; ORGANISM: Aspergillus fumigatus  
US-10-128-714-5164

Query Match 55.9%; Score 21.8; DB 5; Length 4505;  
Best Local Similarity 78.8%; Pred. No. 21;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 3 ACATCACGCGTCTTATAGGAGCGTCTTAAT 35  
||||| ||||| ||||| ||||| ||||| |||||  
Db 2471 ACACGCAACGCGTCTTATAGGAGCGATTAAT 2503

RESULT 10  
US-10-027-632-120434/c  
;; Sequence 120434, Application US/10027632  
;; Publication No. US20020198371A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 120434  
;; LENGTH: 852  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-120434

Query Match 52.8%; Score 20.6; DB 5; Length 852;  
Best Local Similarity 74.3%; Pred. No. 55;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 5 ACTCACGCGTCTTATAGGAGCGTCTTAATCTTA 39  
||||| ||||| ||||| ||||| ||||| |||||  
Db 50 AATCATACTCGATTATATGGAGACTCTTAATCTTA 16

RESULT 11  
US-10-027-632-120435/c  
;; Sequence 120435, Application US/10027632  
;; Publication No. US20020198371A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 120435  
;; LENGTH: 852  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-120435

Query Match 52.8%; Score 20.6; DB 5; Length 852;  
Best Local Similarity 74.3%; Pred. No. 55;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 5 ACTCACGCGTCTTATAGGAGCGTCTTAATCTTA 39  
||||| ||||| ||||| ||||| ||||| |||||  
Db 50 AATCATACTCGATTATATGGAGACTCTTAATCTTA 16

RESULT 12  
US-10-027-632-120434/c  
;; Sequence 120434, Application US/10027632  
;; Publication No. US20030204075A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 120434  
;; LENGTH: 852  
;; TYPE: DNA  
;; ORGANISM: Human



Best Local Similarity 75.8%; Pred. No. 70;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 ACACACACACCGCTCTTATAGCGAGCGTCTTAAT 35  
Db 52 ACACACACACCGCTCTTATAGCGAGCGTCTTAAT 84

Search completed: January 30, 2006, 05:23:21  
Job time : 349 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:26:26 ; Search time 392.5 Seconds  
(without alignments)  
82.538 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4

Perfect score: 39  
Sequence: 1 gaacactacacgcgtttatagcgcgtctaattcta 39

Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.New.\*

- 1: /cgm2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 2: /cgm2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /cgm2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /cgm2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 5: /cgm2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq.\*
- 6: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 8: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 9: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 10: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 11: /cgm2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	51.3	953	7	US-10-750-185-47026 Sequence 47026, A
2	20	51.3	953	7	US-10-750-623-47026 Sequence 47026, A
C 3	19.6	50.3	1036	7	US-10-750-185-53947 Sequence 53947, A
C 4	19.6	50.3	1036	7	US-10-750-623-53947 Sequence 53947, A
C 5	19.4	49.7	2424	7	US-10-750-185-26886 Sequence 26886, A
C 6	19.4	49.7	2424	7	US-10-750-623-26886 Sequence 26886, A
C 7	19.4	49.7	3518	8	US-11-136-527-485 Sequence 485, App
8	19	48.7	732	7	US-10-750-185-53265 Sequence 53265, A
9	19	48.7	732	7	US-10-750-623-53265 Sequence 53265, A
10	18.8	48.2	201	7	US-10-995-561-27542 Sequence 27542, A
11	18.8	48.2	201	7	US-10-995-561-27569 Sequence 27569, A
12	18.8	48.2	537	7	US-10-793-626-1175 Sequence 1175, App
13	18.8	48.2	648	7	US-10-750-185-42748 Sequence 42748, A
14	18.8	48.2	648	7	US-10-750-623-42748 Sequence 42748, A
C 15	18.8	48.2	2226	7	US-10-793-626-1177 Sequence 1177, App
16	18.8	48.2	3511	7	US-10-793-626-1177 Sequence 1177, App
17	18.8	48.2	3705	7	US-10-793-626-3592 Sequence 3592, App
18	18.8	48.2	43436	7	US-10-995-561-13240 Sequence 13240, A
C 19	18.6	47.7	1738	7	US-10-750-185-49884 Sequence 49884, A
C 20	18.6	47.7	1738	7	US-10-750-623-49884 Sequence 49884, A
C 21	18.6	47.7	1741	7	US-10-750-185-34115 Sequence 34115, A
C 22	18.6	47.7	1741	7	US-10-750-623-34115 Sequence 34115, A

ALIGNMENTS

RESULT 1

US-10-750-185-47026  
; Sequence 47026, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 47026  
; LENGTH: 953  
; TYPE: DNA  
; ORGANISM: Bovine 19866880944977

Query Match 51.3%; Score 20; DB 7; Length 953;  
Best Local Similarity 72.2%; Pred. No. 7.2;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 ACATCTACACGCGTCTATAGCGAGCGTCTAATCT 38  
DB 133 ATATCTACATACATCTTATATCTCAATGCTAATCT 168

RESULT 2

US-10-750-623-47026  
; Sequence 47026, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom

Sequence 6622, Ap  
Sequence 2526, Ap  
Sequence 51151, A  
Sequence 51151, A  
Sequence 59, Appl  
Sequence 58, Appl  
Sequence 5, Appl  
Sequence 47157, A  
Sequence 47157, A  
Sequence 58552, A  
Sequence 58552, A  
Sequence 43979, A  
Sequence 43979, A  
Sequence 43979, A  
Sequence 13431, A  
Sequence 13431, A  
Sequence 7552, Ap  
Sequence 26047, A  
Sequence 26047, A  
Sequence 71, Appl  
Sequence 60653, A  
Sequence 60653, A  
Sequence 3456, Ap  
Sequence 13474, A



; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26886  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Bovine 19866880518817  
US-10-750-623-26886

Query Match 49.7%; Score 19.4; DB 7; Length 2424;  
Best Local Similarity 70.3%; Pred. No. 16;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GAACACTCACAGCGCTTATAGCGAGCGTCTTAATTC 37  
Db 621 GCACACTCACACTCATTTGATGAGCGGTTTATTC 585

## RESULT 7

US-11-136-527-485/c  
; Sequence 485, Application US/11136527  
; Publication No. US20050287570A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William M  
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
; FILE REFERENCE: 031896-041000 (AM101086)  
; CURRENT APPLICATION NUMBER: US/11/136,527  
; CURRENT FILING DATE: 2005-05-25  
; PRIOR APPLICATION NUMBER: US 60/574,294  
; PRIOR FILING DATE: 2005-05-26  
; NUMBER OF SEQ ID NOS: 362830  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 485  
; LENGTH: 3518  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus

US-11-136-527-485

Query Match 49.7%; Score 19.4; DB 8; Length 3518;  
Best Local Similarity 70.3%; Pred. No. 17;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3 ACATCTCACAGCGCTTATAGCGAGCGTCTTAATTC 39  
Db 786 ACCCTCACAGTGCCTTATAGCGCACCCCTGATGTTA 750

## RESULT 8

US-10-750-185-53265  
; Sequence 53265, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 53265  
; LENGTH: 732  
; TYPE: DNA  
; ORGANISM: Bovine 19866881011834

US-10-750-185-53265

Query Match 48.7%; Score 19; DB 7; Length 732;  
Best Local Similarity 71.4%; Pred. No. 20;  
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 AACACTCACAGCGCTTATAGCGAGCGTCTTAATT 36  
Db 259 AACACTCATTACATATATTTTAAGCTTCTTAATT 293

## RESULT 9

US-10-750-623-53265  
; Sequence 53265, Application US/10750623  
; Publication No. US20050287531A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-1  
; CURRENT APPLICATION NUMBER: US/10/750,623  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 53265  
; LENGTH: 732  
; TYPE: DNA  
; ORGANISM: Bovine 19866881011834

US-10-750-623-53265

Query Match 48.7%; Score 19; DB 7; Length 732;  
Best Local Similarity 71.4%; Pred. No. 20;  
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 AACACTCACAGCGCTTATAGCGAGCGTCTTAATT 36  
Db 259 AACACTCATTACATATATTTTAAGCTTCTTAATT 293

## RESULT 10

US-10-995-561-27542  
; Sequence 27542, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27542  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-995-561-27542

Query Match 48.2%; Score 18.8; DB 7; Length 201;  
Best Local Similarity 68.4%; Pred. No. 20;  
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 AACACTCACAGCGCTTATAGCGAGCGTCTTAATCTA 39  
Db 29 AACCTCCCATGCTCTCCAGAACTGAGCGCTCAATCTA 66

```
RESULT 11
US-10-995-561-27569
; Sequence 27569, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27569
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-27569

Query Match      48.2%; Score 18.8; DB 7; Length 201;
Best Local Similarity 68.4%; Pred. No. 20;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy  2 AACACTCACACGCGTCTTATAGCGGCGTCTAAATTCTA 39
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  18 AACCTCCCATGCTTCCAGAACTGAGCGCTCAATTCTA 55

RESULT 12
US-10-793-626-1175
; Sequence 1175, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1175

Query Match      48.2%; Score 18.8; DB 7; Length 537;
Best Local Similarity 76.7%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  9 ACACGCGTCTTATAGCGGCGTCTAAATTCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db  375 ATACGCTTGTAAAGCGTGCCTCAATTCT 404

RESULT 13
US-10-750-185-42748
; Sequence 42748, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1175
```

```
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42748
; LENGTH: 648
; TYPE: DNA
; ORGANISM: Bovine 19866881287324
US-10-750-185-42748

Query Match      48.2%; Score 18.8; DB 7; Length 648;
Best Local Similarity 68.4%; Pred. No. 24;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy  1 GAACACTCACACGCGTCTTATAGCGGCGTCTAAATTCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db  246 GAACATAAAGGGGTCTGCGAATGAGCGTCTAAAGCT 283

RESULT 14
US-10-750-623-42748
; Sequence 42748, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42748
; LENGTH: 648
; TYPE: DNA
; ORGANISM: Bovine 19866881287324
US-10-750-623-42748

Query Match      48.2%; Score 18.8; DB 7; Length 648;
Best Local Similarity 68.4%; Pred. No. 24;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy  1 GAACACTCACACGCGTCTTATAGCGGCGTCTAAATTCT 38
    ||||| ||||| ||||| ||||| ||||| |||||
Db  246 GAACATAAAGGGGTCTGCGAATGAGCGTCTAAAGCT 283

RESULT 15
US-10-793-626-1177/c
; Sequence 1177, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1175
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1175
```



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1177
; LENGTH: 2226
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1177

Query Match      48.2%; Score 18.8; DB 7; Length 2226;
Best Local Similarity 76.7%; Pred No. 30;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      9 ACACGCGTCTTATAGCGAGCGTCTAATTCT 38
Db      996 ATACGCTTGTGTTAAGCGTGGTCCAATTCT 967

Search completed: January 30, 2006, 05:36:32
Job time : 392.5 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:23:56 ; Search time 219 Seconds  
(without alignments)  
1186.863 Million cell updates/sec

Title: US-10-663-033-3-REVCOMP-4

Perfect score: 39  
Sequence: 1 gaacactcacgcgtttatagcagcgcttaattcta 39

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	22	56.4	934	2	AAH14231
c 2	22	56.4	1119	3	AAH09495 H. pylori
c 3	21.8	55.9	2295	8	ABT18994 Aspergill
c 4	21.8	55.9	2450	8	ABT18400 Aspergill
c 5	21.8	55.9	2505	8	ABT20216 Aspergill
c 6	21.8	55.9	2505	8	ABT20814 Aspergill
c 7	21.8	55.9	4450	8	ABT17806 Aspergill
c 8	21.8	55.9	4505	8	ABT19620 Aspergill
c 9	20.2	51.8	2454	8	ACA48494 Prokaryot
c 10	20.2	51.8	29773	4	ACA29814 Human cyt
c 11	20.2	51.8	29773	12	ADJ12408 DNA fragm
c 12	20	51.3	20	3	AAH09497 Primer GW
c 13	20	51.3	117754	11	ACN43866 Human gen
c 14	19.8	50.8	2349	13	ADT05658 Haemophil
c 15	19.8	50.8	3469	13	ADT05413 Haemophil
c 16	19.8	50.8	191996	13	ADT05647 Haemophil
c 17	19.8	50.8	263852	13	ADS99460 Murine kl
c 18	19.6	50.3	189158	11	ACN44124 Mouse gen
c 19	19.4	49.7	653	2	AAV52043 Helicobac

c 20	19.4	49.7	685	5	AAH81594
c 21	19.4	49.7	2450	11	ADW22220 Rat hepat
c 22	19.4	49.7	3576	10	ADB58911 Toxicity-
c 23	19.4	49.7	3576	10	ADB53633 Primary r
c 24	19.4	49.7	3576	13	ADV41650 Rat cardi
c 25	19.4	49.7	5758	6	ABL35076 Murine CD
c 26	19.4	49.7	35658	9	ADA02867 Mouse sel
c 27	19.4	49.7	35658	10	ADB72605 Mouse Sel
c 28	19.4	49.7	35658	10	ADC85346 Human itp
c 29	19.4	49.7	35658	12	ADM74462 Murine ca
c 30	19.4	49.7	38630	14	ADZ12686 Murine ca
c 31	19.2	49.2	505	10	ADB50969 Primary r
c 32	19.2	49.2	667	13	ADS61083 Bacterial
c 33	19.2	49.2	2000	6	ABZ17024 Arabidops
c 34	19.2	49.2	2664	11	ACH98752 Klebssteil
c 35	19	48.7	19	3	AAA09496 Primer GW
c 36	19	48.7	2363	4	ABL23910 Drosophil
c 37	19	48.7	2670	11	ACH94275 Klebssteil
c 38	19	48.7	2773	4	ABL04300 Drosophil
c 39	19	48.7	3319	4	ABL04930 Drosophil
c 40	19	48.7	3633	4	ABL04926 Drosophil
c 41	19	48.7	110000	14	AEA42401_28 Continnuation (29 o
c 42	19	48.7	124884	5	AAH74201 Nucleotid
c 43	19	48.7	124884	10	ADA14878 Human her
c 44	19	48.7	124884	11	ADL99489 Varicella
c 45	19	48.7	125157	5	AAH74202 Nucleotid

#### ALIGNMENTS

#### RESULT 1

AAH14231/c

ID AAX14231 standard; DNA; 934 BP.

XX AC AAX14231;

XX DT 31-MAR-1999 (first entry)

XX DB H. pylori GHPO 1270 gene.

XX KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

XX KW peptic ulcer disease; ss.

XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers

XX FT CDS 51..884

XX FT /\*tag= a

XX FN WO9843478-A1.

XX PD 08-OCT-1998.

XX PF 01-APR-1998; 98WO-US006371.

XX PR 01-APR-1997; 97US-00833457.

XX PR 24-JUN-1997; 97US-00881227.

XX PR 29-JUL-1997; 97US-00902615.

XX PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;

XX DR WPI; 1998-542293/46.

XX DR P-ESDB; AAW98512.

XX PT New isolated Helicobacter polynucleotides - used to develop products for

XX PT the diagnosis, prevention and treatment of Helicobacter infections and

XX PT gastrointestinal diseases.

XX PS Claim 1; Page 964-966; 2054pp; English.

XX This sequence represents a polynucleotide of the invention. It was  
 CC isolated from *Helicobacter pylori* and encodes a H.pylori GHPO protein.  
 CC The polypeptides can be used for preventing or treating *Helicobacter*  
 CC infections, and gastroduodenal diseases associated with these infections,  
 CC including acute, chronic, and atrophic gastritis, and peptic ulcer  
 CC diseases, e.g. gastric and duodenal ulcers. They can also be used for the  
 CC production of antibodies. The products can also be used for detection and  
 CC diagnosis

XX SQ Sequence 934 BP; 306 A; 182 C; 181 G; 265 T; 0 U; 0 Other;

Query Match 56.4%; Score 22; DB 2; Length 934;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39  
 Db 844 TTATAGCGAGCGTCTAATTCTA 823

## RESULT 2

AAA09495  
 ID AAA09495 standard; DNA; 1119 BP.

XX AC AAA09495;

XX DT 29-AUG-2000 (first entry)

XX DE H. pylori UA802 fucT2 coding sequence.

XX KW fucT2; alpha-1,2-fucosyltransferase; Lewis Y; glycoconjugate;  
 KW fucosylated oligosaccharide; tumor associated carbohydrate antigen;  
 KW tumour; 88.

XX OS *Helicobacter pylori*.

XX FH Key Location/Qualifiers  
 FT CDS 137..1039  
 FT /\*tag= a

XX PN WO200026383-A1.

XX PD 11-MAY-2000.

XX DP 03-NOV-1999; 99WO-CA001031.

XX PR 04-NOV-1998; 98US-0107268P.

XX PR 02-NOV-1999; 99US-00433598.

XX PA (UYAL-) UNIV ALBERTA.

XX PI Taylor DE, Wang G, Palcic M;

XX DR WPI; 2000-365628/31.

XX DR P-PSDB; AAY92713.

XX PT *Helicobacter pylori* alpha1,2-fucosyltransferase enzymes useful for  
 PT producing a fucosylated oligosaccharide and for diagnosing malignancies  
 PT related to H. pylori infections.

XX PS Claim 10; Fig 1D; 71pp; English.

XX CC *Helicobacter pylori* fucT2 gene encodes alpha-1,2-fucosyltransferase. The  
 CC UA802 fucT2 contains a single open reading frame encoding a protein of  
 CC 300 amino acids (see AAY92713). It was designated Hp fucT2 to distinguish  
 CC it from the previously identified alpha-1,3-fucT, which was designated  
 CC fucT. Hp fucT2 gene has a unique feature in its centre region. In  
 CC addition to a poly C tract, a sequence of TAA repeats (imperfect)  
 CC immediately following the poly C sequence were identified. The changes of  
 CC the repeat number of both tracts contribute to the variation of the fucT2  
 CC genotype (on or off status) in different strains. The UA802 fucT2 gene is  
 CC intact, whereas in strain 26695 it is frameshifted. However, the 26695

CC strain does produce a full-length protein - it was discovered that the  
 CC gene contains a slippery heptanucleotide sequence identical to that found  
 CC in the mRNA of *Escherichia coli* dnaX. This allows for ribosome slippage  
 CC in translation. The Hp fucT2 polypeptide catalyzes the synthesis of Lewis  
 CC Y (claimed), and other fucosylated oligosaccharides such as Lewis X,  
 CC Lewis B and H type 1. The alpha-1,2-fucosyltransferase enzymes are useful  
 CC for producing fucosylated oligosaccharides such as Lewis X, Lewis Y,  
 CC Lewis B and H type 1, which are structurally similar to certain tumor  
 CC associated carbohydrate antigens found in mammals. These glycoconjugates  
 CC have research and diagnostic utility for the development of assays and  
 CC reagents (e.g. antibodies) for detecting H. pylori and associated  
 CC mammalian tumours

XX SQ Sequence 1119 BP; 363 A; 195 C; 227 G; 334 T; 0 U; 0 Other;

Query Match 56.4%; Score 22; DB 3; Length 1119;  
 Best Local Similarity 100.0%; Pred. No. 7.3;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 TTATAGCGAGCGTCTAATTCTA 39

Db 1098 TTATAGCGAGCGTCTAATTCTA 1119

## RESULT 3

ABT18994

ID ABT18994 standard; DNA; 2295 BP.

XX AC ABT18994;

XX DT 16-APR-2003 (first entry)

XX DE *Aspergillus fumigatus* essential gene #1352.

XX KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
 KW cancer; contamination; biofilm; antibody; immune response; ds.

XX OS *Aspergillus fumigatus*.

XX PN WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Jiang B, Tishkoff D, Zamudio C, Broshkin AM, Hu W, Lemieux SM;

XX DR WPI; 2003-093124/08.

XX PT New purified or isolated nucleic acids of essential genes of *Aspergillus*  
 PT *fumigatus*, useful for treating or preventing infections by A. *fumigatus*,  
 PT or for treating a non-infectious disease in a subject e.g. cancer.

XX PS Disclosure; Page; 175pp; English.

XX CC The invention relates to novel purified or isolated nucleic acids of  
 CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
 CC the invention are used to treat or prevent infections by a pathogenic  
 CC organism such as A. *fumigatus*, to treat a non-infectious disease in a  
 CC subject (e.g. cancer), to prevent or contain contamination of an object  
 CC by A. *fumigatus*, or to prevent or inhibit formation on a surface of a  
 CC biofilm comprising A. *fumigatus*. The polynucleotides are useful for  
 CC expressing recombinant protein for characterisation, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of A.

CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of *Aspergillus fumigatus* of the invention

XX SQ Sequence 2295 BP; 604 A; 607 C; 603 G; 481 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 2295;  
 Best Local Similarity 78.8%; Pred. No. 10;  
 Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 ACACCTCACACGCGTCTTATAGCGAGCGTCTAAT 35

DB 1546 ACACGCAACGCGTCTTATAGCGAGGAGATAT 1578

RESULT 4

ABT18400

ID ABT18400 standard; DNA; 2450 BP.

XX AC

XX AC

DT 16-APR-2003 (first entry)

DE *Aspergillus fumigatus* essential gene #759.

KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
 KW cancer; contamination; biofilm; antibody; immune response; ds.

XX *Aspergillus fumigatus*.

XX WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of *Aspergillus*  
 XX fumigatus, useful for treating or preventing infections by *A. fumigatus*,  
 XX or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of  
 XX essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of  
 XX the invention are used to treat or prevent infections by a pathogenic  
 XX organism such as *A. fumigatus*, to treat a non-infectious disease in a  
 XX subject (e.g. cancer), to prevent or contain contamination of an object  
 XX by *A. fumigatus*, or to prevent or inhibit formation on a surface of a

CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for  
 CC expressing recombinant protein for characterisation, screening or  
 CC therapeutic use, as markers for host tissues in which the pathogenic  
 CC organisms invade or reside, for comparing with the DNA sequence of *A.*  
 CC fumigatus to identify duplicated genes or paralogues having the same or  
 CC similar biochemical activity and/or function, for comparing with DNA  
 CC sequences of other related or distant pathogenic organisms to identify  
 CC potential orthologous essential or virulence genes, for selecting and  
 CC making oligomers for attachment to a nucleic acid array for examination  
 CC of expression patterns, for raising anti-protein antibodies, as an  
 CC antigen to raise anti-DNA antibodies or to elicit another immune  
 CC response, and for identifying polynucleotides encoding the other protein  
 CC with which binding occurs or to identify inhibitors of the binding  
 CC interaction. The polypeptides may be used to raise antibodies or to  
 CC elicit immune response, as a reagent in assays designed to quantitatively  
 CC determine levels of the protein in biological fluids, as a marker for  
 CC host tissues in which pathogenic organism invade or reside, and to  
 CC isolate correlative receptors or ligands in the case of virulence  
 CC factors. This polynucleotide sequence represents one of the essential  
 CC genes of *Aspergillus fumigatus* of the invention

XX SQ Sequence 2450 BP; 644 A; 673 C; 628 G; 505 T; 0 U; 0 Other;

Query Match 55.9%; Score 21.8; DB 8; Length 2450;

Best Local Similarity 78.8%; Pred. No. 10;

Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 ACACCTCACACGCGTCTTATAGCGAGCGTCTAAT 35

DB 1546 ACACGCAACGCGTCTTATAGCGAGGAGATAT 1578

RESULT 5

ABT20216

ID ABT20216 standard; DNA; 2505 BP.

XX AC

XX AC

DT 16-APR-2003 (first entry)

DE *Aspergillus fumigatus* essential gene #2574.

KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;  
 KW cancer; contamination; biofilm; antibody; immune response; ds.

XX *Aspergillus fumigatus*.

XX WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX WPI; 2003-093124/08.

XX New purified or isolated nucleic acids of essential genes of *Aspergillus*  
 XX fumigatus, useful for treating or preventing infections by *A. fumigatus*,  
 XX or for treating a non-infectious disease in a subject e.g. cancer.

XX Disclosure; Page; 175pp; English.

XX The invention relates to novel purified or isolated nucleic acids of  
 XX essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of



PT	New purified or isolated nucleic acids of essential genes of Aspergillus
PT	fumigatus, useful for treating or preventing infections by A. fumigatus,
PT	or for treating a non-infectious disease in a subject e.g. cancer.
XX	
PS	Disclosure; Page; 175pp; English.
XX	
CC	The invention relates to novel purified or isolated nucleic acids of
CC	essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC	the invention are used to treat or prevent infections by a pathogenic
CC	organism such as A. fumigatus, to treat a non-infectious disease in a
CC	subject (e.g. cancer), to prevent or contain contamination of an object
CC	by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC	biofilm comprising A. fumigatus. The polynucleotides are useful for
CC	expressing recombinant protein for characterisation, screening or
CC	therapeutic use, as markers for host tissues in which the pathogenic
CC	organisms invade or reside, for comparing with the DNA sequence of A.
CC	fumigatus to identify duplicated genes or paralogues having the same or
CC	similar biochemical activity and/or function, for comparing with DNA
CC	sequences of other related or distant pathogenic organisms to identify
CC	potential orthologous essential or virulence genes, for selecting and
CC	making oligomers for attachment to a nucleic acid array for examination
CC	of expression patterns, for raising anti-protein antibodies, as an
CC	antigen to raise anti-DNA antibodies or to elicit another immune
CC	response, and for identifying polynucleotides encoding the other protein
CC	with which binding occurs or to identify inhibitors of the binding
CC	interaction. The polypeptides may be used to raise antibodies or to
CC	elicit immune response, as a reagent in assays designed to quantitatively
CC	determine levels of the protein in biological fluids, as a marker for
CC	host tissues in which pathogenic organism invade or reside, and to
CC	isolate correlative receptors or ligands in the case of virulence
CC	factors. This polynucleotide sequence represents one of the essential
CC	genes of Aspergillus fumigatus of the invention
XX	
SQ	Sequence 4450 BP; 1174 A; 1126 C; 1152 G; 998 T; 0 U; 0 Other;
	Query Match            55.9%;   Score 21.8; DB 8; Length 4450;
	Best Local Similarity   78.8%;   Pred. No. 12;
	Matches   26; Conservative   0; Mismatches   7; Indels   0; Gaps   0
Qy	3 ACACCTCACACGCGTCTTAGCGAGCGGTCAAT 35 
Db	2546 ACACGAACAACGCGTTTAGCGAGGAGATAAT 2578 
RESULT 8	
ABT19620	
ID	ABT19620 standard; DNA; 4505 BP.
XX	
AC	ABT19620;
XX	
DT	16-APR-2003 (first entry)
XX	
DE	Aspergillus fumigatus essential gene #1978.
XX	
KW	Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW	cancer; contamination; biofilm; antibody; immune response; ds.
XX	
OS	Aspergillus fumigatus.
XX	
FN	WO200286090-A2.
XX	
PD	31-OCT-2002.
XX	
PP	23-APR-2002; 2002WO-US013142.
XX	
FR	23-APR-2001; 2001US-0285697P.
FR	27-APR-2001; 2001US-0287066P.
PR	05-JUN-2001; 2001US-0295890P.
PR	09-JUL-2001; 2001US-0303899P.
PR	31-AUG-2001; 2001US-0316362P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	

PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU44624.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 36364; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2454 BP; 723 A; 471 C; 523 G; 737 T; 0 U; 0 Other;  
  
Query Match 51.8%; Score 20.2; DB 8; Length 2454;  
Best Local Similarity 88.0%; Pred. No. 57;  
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Oy 2 AACACTACAGCGCTCTATAGCGA 26  
|||||  
Db 1572 AACACTACAGCGCTCAATAGCGA 1548  
  
RESULT 10  
AAS29814  
ID AAS29814 standard; DNA; 29973 BP.  
XX  
XX AAS29814;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human cytoskeletal element-related polypeptide encoding genomic DNA #1.  
XX  
XX Cytoskeletal element-related protein; human; mouse; rabbit; goat; horse;  
KW cat; dog; chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;

KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; food preservative;  
KW tissue regeneration; anti-infertility; food additive.  
XX Homo sapiens.  
XX  
XX WO200155168-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001331.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.



PR	14-SEP-2000;	2000US-202330633P;
PR	14-SEP-2000;	2000US-202330643P;
PR	14-SEP-2000;	2000US-20233065P;
PR	21-SEP-2000;	2000US-02342233P;
PR	21-SEP-2000;	2000US-02342274P;
PR	25-SEP-2000;	2000US-0234997P;
PR	25-SEP-2000;	2000US-0234998P;
PR	26-SEP-2000;	2000US-0235484P;
PR	27-SEP-2000;	2000US-0235834P;
PR	27-SEP-2000;	2000US-0235836P;
PR	29-SEP-2000;	2000US-0236327P;
PR	29-SEP-2000;	2000US-0236368P;
PR	29-SEP-2000;	2000US-0236369P;
PR	29-SEP-2000;	2000US-0236370P;
PR	02-OCT-2000;	2000US-0236802P;
PR	02-OCT-2000;	2000US-0237037P;
PR	02-OCT-2000;	2000US-0237038P;
PR	02-OCT-2000;	2000US-0237039P;
PR	02-OCT-2000;	2000US-0237040P;
PR	13-OCT-2000;	2000US-0239935P;
PR	13-OCT-2000;	2000US-0239937P;
PR	20-OCT-2000;	2000US-0240960P;
PR	20-OCT-2000;	2000US-0241221P;
PR	20-OCT-2000;	2000US-0241785P;
PR	20-OCT-2000;	2000US-0241786P;
PR	20-OCT-2000;	2000US-0241787P;
PR	20-OCT-2000;	2000US-0241808P;
PR	20-OCT-2000;	2000US-0241809P;
PR	20-OCT-2000;	2000US-0241826P;
PR	01-NOV-2000;	2000US-0244617P;
PR	08-NOV-2000;	2000US-0246474P;
PR	08-NOV-2000;	2000US-0246475P;
PR	08-NOV-2000;	2000US-0246476P;
PR	08-NOV-2000;	2000US-0246477P;
PR	08-NOV-2000;	2000US-0246478P;
PR	08-NOV-2000;	2000US-0246523P;
PR	08-NOV-2000;	2000US-0246524P;
PR	08-NOV-2000;	2000US-0246525P;
PR	08-NOV-2000;	2000US-0246526P;
PR	08-NOV-2000;	2000US-0246527P;
PR	08-NOV-2000;	2000US-0246528P;
PR	08-NOV-2000;	2000US-0246532P;
PR	08-NOV-2000;	2000US-0246609P;
PR	08-NOV-2000;	2000US-0246610P;
PR	08-NOV-2000;	2000US-0246611P;
PR	08-NOV-2000;	2000US-0246613P;
PR	17-NOV-2000;	2000US-0249207P;
PR	17-NOV-2000;	2000US-0249208P;
PR	17-NOV-2000;	2000US-0249209P;
PR	17-NOV-2000;	2000US-0249210P;
PR	17-NOV-2000;	2000US-0249211P;
PR	17-NOV-2000;	2000US-0249212P;
PR	17-NOV-2000;	2000US-0249213P;
PR	17-NOV-2000;	2000US-0249214P;
PR	17-NOV-2000;	2000US-0249215P;
PR	17-NOV-2000;	2000US-0249244P;
PR	17-NOV-2000;	2000US-0249245P;
PR	17-NOV-2000;	2000US-0249265P;
PR	17-NOV-2000;	2000US-0249276P;
PR	17-NOV-2000;	2000US-0249299P;
PR	17-NOV-2000;	2000US-0249300P;
PR	01-DEC-2000;	2000US-02501160P;
PR	01-DEC-2000;	2000US-0250391P;
PR	05-DEC-2000;	2000US-0251030P;
PR	05-DEC-2000;	2000US-0251988P;
PR	05-DEC-2000;	2000US-0256719P;
PR	06-DEC-2000;	2000US-0251479P;
PR	06-DEC-2000;	2000US-0251856P;

PR	08-DEC-2000; 2000US-0251868P.
PR	08-DEC-2000; 2000US-0251869P.
PR	08-DEC-2000; 2000US-0251989P.
PR	08-DEC-2000; 2000US-0251990P.
PR	11-DEC-2000; 2000US-0254097P.
PR	05-JAN-2001; 2001US-0259678P.
XX	
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	
FI	Rosen CA, Barash SC, Ruben SM;
XX	
XX	WPI; 2001-476182/51.
DR	
XX	
PT	Novel isolated human cytoskeletal element-related polypeptide useful for
PT	diagnosis/treatment of neoplastic disorders, disorders associated with
PT	neural transmission, chromosomal abnormalities, autoimmune disorders.
XX	
XX	Claim 1; SEQ ID NO 99; 505pp; English.
PS	
XX	
CC	Sequences AAS29814-AAS29838 represent genomic DNA molecules, which encode
CC	the cytoskeletal element-related polypeptides of the invention.
CC	Cytoskeletal polypeptides and their associated polynucleotides are useful
CC	in the diagnosis, treatment and prevention of various types of disorders
CC	in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or
CC	sheep. A pathological condition can be determined by determining the
CC	presence or absence of a mutation in a cytoskeletal polynucleotide. The
CC	treatable disorders include autoimmune diseases such as rheumatoid
CC	arthritis, hyperproliferative disorders such as neoplasms of the breast
CC	or liver, cardiovascular disorders such as cardiac arrest,
CC	cerebrovascular disorders such as cerebral ischaemia, nervous system
CC	disorders such as Alzheimer's disease, infections caused by bacteria,
CC	viruses and fungi, ocular disorders such as corneal infection, endocrine
CC	disorders such as premature labour and infertility, gastrointestinal
CC	disorders such as Crohn's disease, renal disorders such as
CC	glomerulonephritis and respiratory disorders such as asthma. The
CC	polypeptides can also be used to aid wound healing, to prevent skin aging
CC	due to sunburn, to maintain organs before transplantation, to regenerate
CC	tissues and in chemotaxis. The polypeptides can also be used as a food
CC	additive or preservative to increase or decrease storage capabilities.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at <a href="http://ftp.wipo.int/pub/published/pct/sequences">ftp.wipo.int/pub/published/pct/sequences</a>

KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;  
 KW antipsoriatic; antibacterial; osteopathic; dermatological; antigout;  
 KW immunomodulator; antiarrhythmic; cardiac; neotropic; antileptic;  
 KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;  
 KW antidiabetic; anabolic; hypertensive; vulnery; ds.

OS Homo sapiens.

XX US2004010132-A1.

XX 15-JAN-2004.

XX 30-OCT-2001; 2001US-00984429.

XX 09-OCT-1997; 97US-0061463P.

XX 09-OCT-1997; 97US-0061527P.

XX 09-OCT-1997; 97US-0061529P.

XX 09-OCT-1997; 97US-0061532P.

XX 09-OCT-1997; 97US-0061536P.

XX 09-OCT-1997; 97US-0071498P.

XX 08-OCT-1998; 98WO-US021142.

XX 08-APR-1999; 99US-00288143.

XX 01-NOV-2000; 2000US-0244591P.

XX (ROSE/) ROSEN C A.

XX (BREW/) BREWER L A.

XX (DUAN/) DUAN R D.

XX (RUBE/) RUBEN S M.

XX (FLOR/) FLORENCE K A.

XX (GREE/) GREENE J M.

XX (YOUN/) YOUNG P E.

XX (FERR/) FERRIE A M.

XX (YUGG/) YU G.

XX (FLOR/) FLORENCE C.

XX (EBNE/) EBNER R.

XX (OLSE/) OLSEN H.

XX Rosen CA, Brewer LA, Duan RD, Ruben SM, Florence KA, Greene JM;

XX Young PE, Ferrie AM, Yu G, Florence C, Ebner R, Olsen H;

XX WPI; 2004-090518/09.

XX New isolated nucleic acids and polypeptides, useful for diagnosing,

XX treating, preventing or ameliorating diseases or disorders e.g. cancer,

XX anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's

XX disease.

XX Disclosure; SEQ ID NO 262; 286pp; English.

XX This invention relates to novel polynucleotides encoding human secreted

XX proteins. Specifically, it refers to the vectors, host cells, recombinant

XX and synthetic methods for producing human polynucleotides, polypeptides

XX and antibodies. Furthermore, it relates to screening methods to identify

XX agonists and antagonists that can be used to inhibit or enhance the

XX production and function of the secreted proteins. The present invention

XX describes these compositions as useful for diagnosing, treating or

XX preventing disorders such as cancer, haematopoietic diseases including

XX anemia and multiple myeloma, reproductive system disorders including

XX prostatic and inguinal hernia, musculoskeletal diseases including

XX systemic lupus erythematosus and gout, cardiovascular diseases including

XX arrhythmia and hypernatraemia, mixed fetal diseases including fetal

XX alcohol syndrome and Down's syndrome, excretory diseases including

XX urinary incontinence and renal disorders, neural or sensory disease

XX including Alzheimer's disease and meningitis, respiratory disease

XX including emphysema and occupational lung disease, endocrine diseases

XX including diabetes and glomerulonephritis, digestive diseases including

XX portal hypertension and irritable bowel syndrome and connective tissue or

XX epithelial diseases including scleroderma and epidermolysis bullosa. As

XX such, there are various activities such as cytostatic, antianemic,

XX antiarthritic, antisthmatic, anti-HIV, immunosuppressive,

XX antiinflammatory, antipsoriatic, antibacterial, osteopathic,

XX dermatological, antigout, immunomodulator, antiarrhythmic, cardiac,

XX neotropic, antileptic, nephrotropic, uropathic, neuroprotective,

CC antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and  
 CC vulnery. This polynucleotide is a DNA fragment of a BAC clone that  
 CC encodes a human secreted protein of the invention. NOTE: This sequence  
 CC does not appear in the printed specification but has been obtained in  
 CC electronic format from the US patent office at the following web site  
 CC www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.

XX SQ Sequence 29973 BP; 7911 A; 6282 C; 7047 G; 8733 T; 0 U; 0 Other;

XX Query Match 51.8%; Score 20.2; DB 12; Length 29973;

XX Best Local Similarity 75.8%; Pred. No. 91;

XX Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

XX 6 CTCACACGCGTCTTATAGCGAGCGTCTAAATCT 38

XX 4662 CTCAGCCGTGCTCTTCTGCGACGCTCTAACTCT 4630

XX Db

XX RESULT 12

XX AAA09497/c

XX ID AAA09497 standard; DNA; 20 BP.

XX AC AAA09497;

XX XX

XX 29-AUG-2000 (first entry)

XX DT Primer GW32 for Hp fuct2 gene amplification.

XX DE

XX KW Primer; fuct2; alpha-1,2-fucosyltransferase; Lewis Y; glycoconjugate;

XX KW fucosylated oligosaccharide; tumor associated carbohydrate antigen;

XX KW tumour; ss.

XX OS Helicobacter pylori.

XX XX

XX WO2000026383-A1.

XX PD 11-MAY-2000.

XX XX

XX 03-NOV-1999; 99WO-CA001031.

XX PF

XX 04-NOV-1998; 98US-0107268P.

XX PR 02-NOV-1999; 99US-00433598.

XX XX

XX (UYAL-) UNIV ALBERTA.

XX PA Taylor DE, Wang G, Palcic M;

XX PI WPI; 2000-365628/31.

XX DR

XX Helicobacter pylori alpha1,2-fucosyltransferase enzymes useful for

XX producing a fucosylated oligosaccharide and for diagnosing malignancies

XX related to H. pylori infections.

XX PS

XX Example 1; Page 30; 71pp; English.

XX CC

XX AAA09496-97 are primers corresponding to positions 99980-99962 and 98855-

XX 98874, respectively, in the published Helicobacter pylori genome. They

XX were used to PCR amplify a 1.12 kb fragment containing the H. pylori

XX fuct2 coding sequence (encoding alpha-1,2-fucosyltransferase) from

XX strains 26695 and UA802. The UA802 fuct2 (see AAA09495) contains a single

XX open reading frame encoding a protein of 300 amino acids (see AAY92713).

XX It was designated Hp fuct2 to distinguish it from the previously

XX identified alpha-1,3-fuct, which was designated fuct. Hp fuct2 gene has a

XX unique feature in its centre region. In addition to a poly C tract, a

XX sequence of TAA repeats (imperfect) immediately following the poly C

XX sequence were identified. The changes of the repeat number of both tracts

XX contribute to the variation of the fuct2 genotype (on or off status) in

XX different strains. The Hp fuct2 polypeptide catalyzes the synthesis of

XX Lewis X (claimed), and other fucosylated oligosaccharides such as Lewis

XX X, Lewis B and H type 1. The alpha-1,2-fucosyltransferase enzymes are

XX useful for producing fucosylated oligosaccharides such as Lewis X, Lewis

XX Y, Lewis B and H type 1, which are structurally similar to certain tumor

XX associated carbohydrate antigens found in mammals. These glycoconjugates

CC have research and diagnostic utility for the development of assays and  
CC reagents (e.g. antibodies) for detecting H. pylori and associated  
CC mammalian tumours

XX Sequence 20 BP; 6 A; 4 C; 4 G; 6 T; 0 U; 0 Other;

Query Match 51.3%; Score 20; DB 3; Length 20;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ATAGCGAGCGTCTTAATTCTA 39

Db 20 ATAGCGAGCGTCTTAATTCTA 1

RESULT 13

ACN43866/c

ID ACN43866 standard; DNA; 117754 BP.

AC ACN43866;

XX 18-NOV-2004 (first entry)

XX Human genomic sequence HCG37475.

DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

PI Morris DW;

DR WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
PT Claim 1; SEQ ID NO 28; Opp; English.  
XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX Sequence 117754 BP; 36841 A; 19567 C; 21519 G; 39647 T; 0 U; 180 Other;

Query Match 51.3%; Score 20; DB 11; Length 117754;

Best Local Similarity 72.2%; Pred. No. 1.5e+02;

Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 ACACACACACATCTTAGAGGGATTTCCTAATTCT 38

Db 35011 ACACACACACATCTTAGAGGGATTTCCTAATTCT 34976

RESULT 14

ADT05658/c

XX ADT05658 standard; DNA; 2349 BP.

XX AC ADT05658;

XX 02-DEC-2004 (first entry)

XX Haemophilus influenzae (NTHi) protein coding sequence - SEQ ID 694.  
DE middle ear bacterial infection; nasopharynx bacterial infection; gene;  
XX ds.

XX Haemophilus influenzae.

XX WO2004078949-A2.

XX 16-SEP-2004.

XX 05-MAR-2004; 2004WO-US007001.

XX 06-MAR-2003; 2003US-0453134P.

XX (CHIL-) CHILDRENS HOSPITAL INC.

XX Bakaletz LO, Munson RS, Dyer DW;

XX WPI; 2004-662422/64.

XX P-FSDB; ADT05659.

XX New polynucleotides of nontypeable strain of Haemophilus influenzae,  
PT useful for treating or preventing NTHi bacterial infections of the middle  
PT ear and/or nasopharynx.  
XX Claim 1; SEQ ID NO 694; 88pp; English.  
XX The invention comprises nucleotide sequences (genes) from the genome of a  
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA  
CC sequences of the invention are useful for treating or preventing NTHi  
CC bacterial infections of the middle ear and/or nasopharynx. The present  
CC DNA sequence represents an NTHi protein coding sequence of the invention.  
XX Sequence 2349 BP; 741 A; 431 C; 527 G; 650 T; 0 U; 0 Other;

Query Match 50.8%; Score 19.8; DB 13; Length 2349;

Best Local Similarity 77.4%; Pred. No. 87;

Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 8 CACACGCGTCTTATAGCGAGCGTCTAATTCT 38

Db 1233 CACAAGCGTAGAATAACGAGCGCGTAAATTCT 1203

RESULT 15

ADT05413

ID ADT05413 standard; DNA; 3469 BP.

XX AC ADT05413;

XX 02-DEC-2004 (first entry)

XX Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 449.  
DE middle ear bacterial infection; nasopharynx bacterial infection; ds;  
XX contig.

XX Haemophilus influenzae.

XX WO2004078949-A2.

XX 16-SEP-2004.

PP	05-MAR-2004; 2004WO-US007001.
XX	
PR	06-MAR-2003; 2003US-0453134P.
XX	
PA	(CHIL-) CHILDRENS HOSPITAL INC.
XX	
PI	Bakaletz LO, Munson RS, Dyer DW;
XX	
DR	WPI; 2004-662422/64.
XX	
PT	New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT	useful for treating or preventing NTHi bacterial infections of the middle
PT	ear and/or nasopharynx.
XX	
PS	Example 1; SEQ ID NO 449; 88pp; English.
XX	
CC	The invention comprises nucleotide sequences (genes) from the genome of a
CC	nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC	sequences of the invention are useful for treating or preventing NTHi
CC	bacterial infections of the middle ear and/or nasopharynx. The present
CC	nucleic acid represents an NTHi contig sequence of the invention.
XX	
SQ	Sequence 3469 BP; 962 A; 667 C; 1079 T; 0 U; 0 Other;
	Query Match            50.8%; Score 19.8; DB 13; Length 3469;
	Best Local Similarity    77.4%; Pred. No. 93;
	Matches    24; Conservative    0; Mismatches    7; Indels    0; Gaps    0;
Qy	8 CACAGCGCTCTTATAGCGAGCGCTTAATTCT 38
Db	2149 CACAAGCGGTAGAATAACGAGCGCGCTAATTCT 2179

Search completed: January 30, 2006, 02:04:23  
Job time : 222 secs









Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugiyama, A., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1..1284  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J023087M06"  
 ORIGIN  
 Query Match 57.4%; Score 22.4; DB 15; Length 1284;  
 Best Local Similarity 81.2%; Pred. No. 48;  
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 4 GCGAGCGTCTAAATCTAGAACACTCACACGCG 35  
 |||||  
 Db 329 GCTATCTGCTAAATCTAGAACACTCACACGCG 360  
 |||||  
 RESULT 5  
 AK065982  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:J013045J05, full insert sequence.  
 AK065982  
 VERSION  
 FLI\_CDNA; CAP trapper.  
 KEYWORDS  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
 REFERENCE  
 1  
 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi, S., Sato, K., Kishimoto, N., Yazaki, J., Kishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Otsu, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Miura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yashimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oota, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)  
 12869764  
 2 (bases 1 to 3491)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Kishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,

Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oota, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
 Direct Submission  
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@ias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cDNA/  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1..3491  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J013045J05"  
 FEATURES  
 source  
 Query Match 57.4%; Score 22.4; DB 15; Length 3491;  
 Best Local Similarity 81.2%; Pred. No. 50;  
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 4 GCGAGCGTCTAAATCTAGAACACTCACACGCG 35  
 |||||  
 Db 837 GCTATCTGCTAAATCTAGAACACTCACACGCG 868  
 |||||  
 RESULT 6  
 AE000531/c  
 LOCUS  
 DEFINITION  
 Helicobacter pylori 26695 section 9 of 134 of the complete genome.  
 ACCESSION  
 AE000531 AE000511  
 VERSION  
 AE000531.1 GI:2313173  
 KEYWORDS  
 Helicobacter pylori 26695  
 Helicobacter pylori 26695  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.



REFERENCE AUTHORS	1 (bases 1 to 10530) Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Uterback, T.R., Peterson, J.D., Kelley, J.M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.
TITLE	The complete genome sequence of the gastric pathogen <i>Helicobacter pylori</i>
JOURNAL	Nature 388 (6642), 539-547 (1997)
PUBMED	9252185
REFERENCE AUTHORS	2 (bases 1 to 10530) Tomb, J.-F., White, O., Kerlavage, A.R., Clayton, R.A., Sutton, G.G., Fleischmann, R.D., Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzgerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Uterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. and Venter, J.C.
TITLE	Direct Submission
JOURNAL	Submitted (06-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE AUTHORS	3 (bases 1 to 10530) White, O.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAR-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	source 1..10530 /organism="Helicobacter pylori 26695" /mol_type="genomic DNA" /strain="26695" /db_xref="taxon:85962" 97..185 /gene="trRNA-Ser-3" 97..185 /gene="trRNA-Ser-3" /product="trRNA-Ser" /note="codon recognized: AGC" 394..1227 /gene="HP0091" 394..1227 /gene="HP0091" /note="similar to GB:D13968 SP:P34719 PID:303629 percent identity: 50.74; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="type II restriction enzyme R protein (hsdR)" /protein_id="AAD07162.1" /db_xref="GI:2313174"
gene	/translation="MKNHLPDFIFLKSLSKTSNRTDFDFTMQKCLRNKIKSIANHLNFKLGKDTKSLFKPKYKAFNVNLIIVARDKQIVLDANGNFPYLSYFEDGKVVETRTGLERIICFNRIKOLNDFVCGIIVGLDSNARKNSGKVMENHLSGLFTNAOLNFKVEIRSEFELCOAFQGDIDKDFVCGDKTYPIEAFNFTYISGSKNEVARSYQDLAKFEAFNYEFIMTIDGIGMDAKNKLQEAYSVEIINLSNVNDFISKAQK W"
gene	1221..2054 /gene="HP0092"
CDS	1221..2054 /gene="HP0092" /note="similar to GB:D13968 SP:P34721 PID:303630 percent identity: 55.30; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="type II restriction enzyme M protein (hsdM)" /protein_id="AAD07163.1" /db_xref="GI:2313175"
gene	/translation="WVIAHNSNEIARPIFKSQDQLFTLYQGDNEVLPQFENQDLPILPDPYPLNDGSLGSIQSGKIVSNKGDWKBGIDBNFYQWNNAKKALQDTSGLLI SGTYHNTFSLGCVLQKLDPIKLNIITWQTNPPNPFSCRYLTHSAEQIIVARKSRKHK HGFNYEVLKINNDKQMRDWSPPAIAPWEKVGKHPQKPELALIVRLIMASDENSL IGDPFGSGSTTGIAANLLKRFBIKIEKESBFIKISMDRKIELDARYKEIRSKINDLNH Q"
gene	complement(2074..2514) /gene="HP0093"
CDS	complement(2074..2514) /gene="HP0093" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=1 /product="H. pylori predicted coding region HP0093" /protein_id="AAD07169.2" /db_xref="GI:4580425"
gene	/translation="MILAAKNSVFIHRRGDYVGICGLGIDYQKKALEYMAKRVPMN ELVFVCSLFTQNLGLGYPMDMTTRDKBEAYWMLLMOSQCHQITANSTYSWAA YLIENPKIIGPKHLLGFPHENILCKEHWVKIESHPFVKSKYNA"
gene	complement(2511..2978) /gene="HP0094"
CDS	complement(2511..2978) /gene="HP0094" /note="similar to PID:1197639 PID:1197650 percent identity: 27.62; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="conserved hypothetical protein" /protein_id="AAD07164.1" /db_xref="GI:2313176"
gene	/translation="MAFKVQICGGLGNQMFQYAFKSLQKHSNTPVLLDITSPDMSD RKQLELFPIDLPVSAKAEIAIAKQHLPKLVRDALCMGDFRVSQETVPVSEPKLLK PSRLTYFGYFQDPYFDALSPILKQFTLPPPPKIRIIRIIKKRKNISASFL"
gene	complement(3077..3607) /gene="HP0095"
CDS	complement(3077..3607) /gene="HP0095" /note="hypothetical protein; identified by GeneMark; putative" /codon_start=1 /transl_table=1 /product="H. pylori predicted coding region HP0095" /protein_id="AAD07170.1" /db_xref="GI:2313182"
gene	/translation="MPKPKNTLPCSLSVKMSYPMRLIKWRISLSHKMTLIQILS ILAASASDEBQKCKIDYITRLNAKIASDVYKRVLLVNEICNESLFDKBSV KLSDDLIIQDILVALVDEMLKBDKYQVQHTILKGIIRKRYDEAYSLNSDEDRILLEYQR RLLEHSHASFNSKKFK"
gene	3680..4624 /gene="HP0096"
CDS	3680..4624 /gene="HP0096" /note="similar to GB:L77117 PID:1591676 percent identity: 31.03; identified by sequence similarity; putative" /codon_start=1 /transl_table=1 /product="phosphoglycerate dehydrogenase" /protein_id="AAD07185.1" /db_xref="GI:2313177"
gene	/translation="MKTFFKGVILDAKSVGLKALEVLEKAVADPDFEYVTPPSQIVERS IEABIMVKNKVVITQVLSQPLKLCITATGTDNDVIKSAKALGIEVKNVSAYSTE SVAQHTLACALSILGRINDYRCKSGEYSQSLFTHSIDIKMGLIKGSONGVIGLGT IGKRVAKLAQAPGAKVYVYSPDKKEEYERLSKDLATSDIISIHAFLNSTRDLIA LKELQSGAILINVGKGVINEKLELIELKDLATSDVFKSPFKDHAFLNPK IQNKLTLTPHIAWAYSLSKTLVETKENTIQDFLASQK"
gene	complement(4781..5494) /gene="HP0097"
CDS	complement(4781..5494) /gene="HP0097" /note="hypothetical protein; identified by GeneMark; putative"

```
/codon_start=1
/transl_table=11
/product="H. pylori predicted coding region HP0097"
/protein_id="AAD07171.1"
/db_xref="GI:2313183"
/translation="MKKIVLVIAIALLMSACASYKITPEHVTSYNNGIQVMTSTQAKSK
VQLBIAQSKLGLNPLVLYAAQVIEGSPVFSRKAISSVINTNLPLVLSLRQVMK
SSDFEGILASITAVPTPIDNVNMITPPMFYGGGFLAYNGMTYGGMYGPGFG
MMMDVDVEQSVQESQALKAINALYNKNTLVNVEKAGGFFVVDTKNTKLTGPGVV
VKVLEDEIHTFKIDISKM"
complement(5599..7059)
/gene="HP0098"
CDS
complement(5599..7059)
/gene="HP0098"
/note="similar to GB:L42023 SP:P44503 PID:1003078
PID:122197 PID:1204345 percent identity: 32.89;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="threonine synthase (thrc)"
/protein_id="AAD07166.1"
/db_xref="GI:2313178"
/translation="MPFVTRSLKEKKIDFIALNPNAPKGGYTLTERPETLQMQDC
LNSYNDLVECFRGLGLETIPKNLLASALKRYENFDNPKNPAPFALNERLFWQELVH
GPSLAFKDMALQPLASFLSNLAVGNKRYLMLVSTGDTGTPATLESAGMPNVFVCH
YPKDGTSLVOKLQWVOSANLKVFGISGDPDQNALKNLLKDDDFEALKAQCLKL
SVANSVNGRTAFQIVVHMGFLRYLKGAINSEKILTALPSGNFGNALGAFYAKKM
GLNIDKIKVYNSNDVLRFEIETGRTYDLTHRSLKQTSYSPAMDILKSNVERALPSLFG
FERTLELMQALEEKFTALKPELALIQEHFSCASCSDEACLKTIQVYAEHQYLLDP
HTATPALNASLNTKHTLIVSATSYERFPRTLLTALNEQKNDNDKAALETIKNSYNT
DSQRLDPLFERGIRKHQBYLKLNEIKSSILLMLSESH"
7263..9290
/gene="HP0099"
CDS
7263..9290
/gene="HP0099"
/note="similar to GB:L29189 SP:P39216 PID:459689
GB:AL009126 percent identity: 32.81; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="methyl-accepting chemotaxis protein (tlpA)"
/protein_id="AAD07167.1"
/db_xref="GI:2313179"
/translation="MSKGLSIGNKIILCVLIVCVISLGSLSNRSVKILKESALH
SMQDSLHKYNEVQVLENTYTSNGIVKEMLPKDKREIKIGLLKNFILANSHVAGVS
```

Query Match 57.4%; Score 22.4; DB 1; Length 10530;  
Best Local Similarity 95.8%; Pred. No. 51;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	Db	16	TTCTAGAACTACACGCGTCTT	39
		3123	TGCTAGAACTACACGCGTCTT	3100

RESULT 7  
AP008210\_304  
WPCOMMENT

Sequence split into 355 fragments LOCUS AP008210 Accession AP008210

Fragment Name	Begin	End
AP008210_000	1	110000
AP008210_001	100001	210000
AP008210_002	200001	310000
AP008210_003	300001	410000
AP008210_004	400001	510000
AP008210_005	500001	610000
AP008210_006	600001	710000
AP008210_007	700001	810000
AP008210_008	800001	910000
AP008210_009	900001	1010000
AP008210_010	1000001	1110000
AP008210_011	1100001	1210000
AP008210_012	1200001	1310000

AP008210_013	1300001	1410000
AP008210_014	1400001	1510000
AP008210_015	1500001	1610000
AP008210_016	1600001	1710000
AP008210_017	1700001	1810000
AP008210_018	1800001	1910000
AP008210_019	1900001	2010000
AP008210_020	2000001	2110000
AP008210_021	2100001	2210000
AP008210_022	2200001	2310000
AP008210_023	2300001	2410000
AP008210_024	2400001	2510000
AP008210_025	2500001	2610000
AP008210_026	2600001	2710000
AP008210_027	2700001	2810000
AP008210_028	2800001	2910000
AP008210_029	2900001	3010000
AP008210_030	3000001	3110000
AP008210_031	3100001	3210000
AP008210_032	3200001	3310000
AP008210_033	3300001	3410000
AP008210_034	3400001	3510000
AP008210_035	3500001	3610000
AP008210_036	3600001	3710000
AP008210_037	3700001	3810000
AP008210_038	3800001	3910000
AP008210_039	3900001	4010000
AP008210_040	4000001	4110000
AP008210_041	4100001	4210000
AP008210_042	4200001	4310000
AP008210_043	4300001	4410000
AP008210_044	4400001	4510000
AP008210_045	4500001	4610000
AP008210_046	4600001	4710000
AP008210_047	4700001	4810000
AP008210_048	4800001	4910000
AP008210_049	4900001	5010000
AP008210_050	5000001	5110000
AP008210_051	5100001	5210000
AP008210_052	5200001	5310000
AP008210_053	5300001	5410000
AP008210_054	5400001	5510000
AP008210_055	5500001	5610000
AP008210_056	5600001	5710000
AP008210_057	5700001	5810000
AP008210_058	5800001	5910000
AP008210_059	5900001	6010000
AP008210_060	6000001	6110000
AP008210_061	6100001	6210000
AP008210_062	6200001	6310000
AP008210_063	6300001	6410000
AP008210_064	6400001	6510000
AP008210_065	6500001	6610000
AP008210_066	6600001	6710000
AP008210_067	6700001	6810000
AP008210_068	6800001	6910000
AP008210_069	6900001	7010000
AP008210_070	7000001	7110000
AP008210_071	7100001	7210000
AP008210_072	7200001	7310000
AP008210_073	7300001	7410000
AP008210_074	7400001	7510000
AP008210_075	7500001	7610000
AP008210_076	7600001	7710000
AP008210_077	7700001	7810000
AP008210_078	7800001	7910000
AP008210_079	7900001	8010000
AP008210_080	8000001	8110000
AP008210_081	8100001	8210000
AP008210_082	8200001	8310000
AP008210_083	8300001	8410000
AP008210_084	8400001	8510000
AP008210_085	8500001	8610000

AP008210_086	8600001	8710000
AP008210_087	8700001	8810000
AP008210_088	8800001	8910000
AP008210_089	8900001	9010000
AP008210_090	9000001	9110000
AP008210_091	9100001	9210000
AP008210_092	9200001	9310000
AP008210_093	9300001	9410000
AP008210_094	9400001	9510000
AP008210_095	9500001	9610000
AP008210_096	9600001	9710000
AP008210_097	9700001	9810000
AP008210_098	9800001	9910000
AP008210_099	9900001	10010000
AP008210_100	10000001	10110000
AP008210_101	10100001	10210000
AP008210_102	10200001	10310000
AP008210_103	10300001	10410000
AP008210_104	10400001	10510000
AP008210_105	10500001	10610000
AP008210_106	10600001	10710000
AP008210_107	10700001	10810000
AP008210_108	10800001	10910000
AP008210_109	10900001	11010000
AP008210_110	11000001	11110000
AP008210_111	11100001	11210000
AP008210_112	11200001	11310000
AP008210_113	11300001	11410000
AP008210_114	11400001	11510000
AP008210_115	11500001	11610000
AP008210_116	11600001	11710000
AP008210_117	11700001	11810000
AP008210_118	11800001	11910000
AP008210_119	11900001	12010000
AP008210_120	12000001	12110000
AP008210_121	12100001	12210000
AP008210_122	12200001	12310000
AP008210_123	12300001	12410000
AP008210_124	12400001	12510000
AP008210_125	12500001	12610000
AP008210_126	12600001	12710000
AP008210_127	12700001	12810000
AP008210_128	12800001	12910000
AP008210_129	12900001	13010000
AP008210_130	13000001	13110000
AP008210_131	13100001	13210000
AP008210_132	13200001	13310000
AP008210_133	13300001	13410000
AP008210_134	13400001	13510000
AP008210_135	13500001	13610000
AP008210_136	13600001	13710000
AP008210_137	13700001	13810000
AP008210_138	13800001	13910000
AP008210_139	13900001	14010000
AP008210_140	14000001	14110000
AP008210_141	14100001	14210000
AP008210_142	14200001	14310000
AP008210_143	14300001	14410000
AP008210_144	14400001	14510000
AP008210_145	14500001	14610000
AP008210_146	14600001	14710000
AP008210_147	14700001	14810000
AP008210_148	14800001	14910000
AP008210_149	14900001	15010000
AP008210_150	15000001	15110000
AP008210_151	15100001	15210000
AP008210_152	15200001	15310000
AP008210_153	15300001	15410000
AP008210_154	15400001	15510000
AP008210_155	15500001	15610000
AP008210_156	15600001	15710000
AP008210_157	15700001	15810000
AP008210_158	15800001	15910000
AP008210_159	15900001	16010000
AP008210_160	16000001	16110000
AP008210_161	16100001	16210000
AP008210_162	16200001	16310000
AP008210_163	16300001	16410000
AP008210_164	16400001	16510000
AP008210_165	16500001	16610000
AP008210_166	16600001	16710000
AP008210_167	16700001	16810000
AP008210_168	16800001	16910000
AP008210_169	16900001	17010000
AP008210_170	17000001	17110000
AP008210_171	17100001	17210000
AP008210_172	17200001	17310000
AP008210_173	17300001	17410000
AP008210_174	17400001	17510000
AP008210_175	17500001	17610000
AP008210_176	17600001	17710000
AP008210_177	17700001	17810000
AP008210_178	17800001	17910000
AP008210_179	17900001	18010000
AP008210_180	18000001	18110000
AP008210_181	18100001	18210000
AP008210_182	18200001	18310000
AP008210_183	18300001	18410000
AP008210_184	18400001	18510000
AP008210_185	18500001	18610000
AP008210_186	18600001	18710000
AP008210_187	18700001	18810000
AP008210_188	18800001	18910000
AP008210_189	18900001	19010000
AP008210_190	19000001	19110000
AP008210_191	19100001	19210000
AP008210_192	19200001	19310000
AP008210_193	19300001	19410000
AP008210_194	19400001	19510000
AP008210_195	19500001	19610000
AP008210_196	19600001	19710000

Query Match 57.4%; Score 22.4; DB 15; Length 110000;

Best Local Similarity 81.2%; Pred. No. 55;

Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 GCGAGCGTCTAAATCTAGACACTCACACGCG 35

Db 63034 GCTATCTGCTAAATCTAGACACTCACACGCG 63065

RESULT 8

OSJN00077 OSJN00077 166804 bp DNA linear PLN 16-APR-2005  
LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNba0041A02,  
DEFINITION complete sequence.

ACCESSION AL606638 GI:21740914

VERSION HTG.

KEYWORDS Oryza sativa (japonica cultivar-group)

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1  
Feng, Q., Zhang, Y., Hao, P., Wang, S., Fu, G., Huang, Y., Li, Y., Zhu, J.,  
Liu, Y., Hu, X., Jia, P., Zhang, Y., Zhao, Q., Ying, K., Yu, S., Tang, Y.,  
Weng, Q., Zhang, L., Lu, Y., Mu, J., Lu, Y., Zhang, L. S., Yu, Z., Fan, D.,  
Liu, X., Lu, T., Li, C., Wu, Y., Sun, T., Lei, H., Li, T., Hu, H., Guan, J.,  
Wu, M., Zhang, R., Zhou, B., Chen, Z., Chen, L., Jin, Z., Wang, R.,  
Yin, H., Cai, Z., Ren, S., Lv, G., Gu, W., Zhu, G., Tu, Y., Jia, J.,  
Zhang, Y., Chen, J., Kang, H., Chen, X., Shao, C., Sun, Y., Hu, Q.,  
Zhang, X., Zhang, W., Wang, L., Ding, C., Sheng, H., Gu, J., Chen, S.,  
Ni, L., Zhu, F., Chen, W., Lan, L., Lai, Y., Cheng, Z., Gu, M., Jiang, J.,  
Li, J., Hong, G., Xue, Y. and Han, B.

TITLE Sequence and analysis of rice chromosome 4







OK 73019, USA  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
6 (bases 1 to 196824)  
Fu, Y., Zhou, L. and Roe, B.A.  
Direct Submission  
Submitted (27-FEB-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
On Feb 27, 2003 this sequence version replaced gi:28316621.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
-----

# FEATURES

source  
Location/Qualifiers  
1. 196824  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="129/SvJ"  
/db\_xref="taxon:10090"  
/chromosome="3"  
/clone="mg1-30a13"  
/clone\_lib="Genome Systems mouse ES cell BAC libraries I  
and II"  
-----

## ORIGIN

Query Match 56.9%; Score 22.2; DB 9; Length 196824;  
Best Local Similarity 77.1%; Pred. No. 69;  
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
  
QY 4 GCGAGCGTCTATTCTAGAACACTCACAGCGCTCT 38  
DB 187650 GCTAGCCTCAATTCAGAGATTTCATGCTCT 187616

## RESULT 11

BV317977  
LOCUS BV317977 613 bp DNA linear STS 26-JAN-2005  
DEFINITION S236P6501RD12.T0 AlaskanMalamute Canis familiaris STS genomic,  
sequence tagged site.  
ACCESSION BV317977  
VERSION BV317977.1 GI:57516469  
KEYWORDS STS.  
SOURCE Canis familiaris (dog)  
ORGANISM Canis familiaris  
Mammalia; Eutheria; Laurasiatheria; Carnivora; Pisipedia; Canidae;  
Canis.  
1 (bases 1 to 613)  
Lindblad-Toh,K.  
The genome sequence of Canis familiaris  
Unpublished (2004)

## REFERENCE

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome  
Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477  
Fax: 6172580903  
Email: kersti@genome.wi.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 613  
Protocol:  
WGS-discovery (WGS):  
Paired-end low-coverage whole genome shotgun reads were generated  
from 9 breeds  
(German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador  
Retriever, English  
Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese  
Water Dog -100,000 each)  
and five other canids (Chinese, Alaskan, Indian and Spanish Gray  
Wolf as well as the

Californian Coyote).  
The WGS reads were placed uniquely on the CanFam1.0 boxer assembly  
and SNP detection was  
carried out by SSAHA-SNP. 863872 reads were annotated as STSs and  
485941 SNPs were  
annotated with alleles from the boxer and the breed or canid from  
which the particular  
read came. The validation rate for these SNPs was estimated at  
approximately 98%.

WGA-discovery (WGA) of Boxer/Poodle SNPs:  
A second set of SNPs was generated using a similar methodology  
except that the contigs  
from the 1.5x poodle assembly (Kirkness 2003) were used instead of  
WGS reads. Since this  
sequence lacked base quality scores, arbitrary quality scores of  
phred 40 were assigned  
before the poodle sequence was placed uniquely on the CanFam1.0  
boxer assembly and SNP  
detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated  
with alleles from the  
boxer and the poodle. The validation rate for these SNPs was  
estimated at approximately TBD%.

Internal-WGA-discovery (I-WGA):  
A third set of SNPs were discovered by comparing reads in the WGA  
assembly. SNPs were  
defined as mismatch positions that had a base quality of >= 30 on  
both reads in a region  
that aligned without gaps, and with at most one additional mismatch  
in the ten flanking  
bases. For each allele, at least one additional read had to confirm  
it. 731476 SNPs were  
annotated with alleles between the two boxer alleles. The  
validation rate for these SNPs  
was estimated at approximately TBD%.

## FEATURES

Location/Qualifiers  
1. 613  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="AlaskanMalamute"  
/db\_xref="taxon:9615"  
/map="20 22-561 45114719-45115265"  
/clone\_lib="AlaskanMalamute"  
<1. >613

## STS

ORIGIN  
Query Match 55.9%; Score 21.8; DB 10; Length 613;  
Best Local Similarity 92.0%; Pred. No. 89;  
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 10 GTCTAATTTCTAGAACACTCACAGC 34  
DB 458 GTCCAAATTTCTAGAACACTCACAC 482

## RESULT 12

AC153673/c  
LOCUS AC153673 84669 bp DNA linear HTG 15-DEC-2004  
DEFINITION Bos taurus clone TAMU-417, WORKING DRAFT SEQUENCE, 5 unordered  
pieces.  
ACCESSION AC153673  
VERSION AC153673.1 GI:56606165  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus

## REFERENCE

AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
Alabrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,







Assembly program: Phrap; version 0.990329  
Consensus quality: 147064 bases at least Q40  
Consensus quality: 157520 bases at least Q30  
Consensus quality: 162179 bases at least Q20  
Insert size: 164635; sum-of-contigs  
Quality coverage: 4.08x in Q20 bases; sum-of-contigs

-----  
NOTE: This is a 'working draft' sequence. It currently consists of 42 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12685 contig of 12685 bp in length  
12786 21378 contig of 8593 bp in length  
21479 31812 contig of 10334 bp in length  
31913 40293 contig of 8381 bp in length  
40394 50693 contig of 10300 bp in length  
50794 57517 contig of 6724 bp in length  
57618 65718 contig of 8101 bp in length  
65819 72757 contig of 6939 bp in length  
72858 80449 contig of 7592 bp in length  
80550 87968 contig of 7419 bp in length  
87969 93179 contig of 5111 bp in length  
93280 99470 contig of 6191 bp in length  
99571 105119 contig of 5549 bp in length  
105220 109327 contig of 4108 bp in length  
109428 114708 contig of 5281 bp in length  
114809 118902 contig of 4094 bp in length  
118903 123406 contig of 4404 bp in length  
123507 126415 contig of 2909 bp in length  
126516 129075 contig of 2560 bp in length  
129176 130072 contig of 897 bp in length  
130173 132075 contig of 1903 bp in length  
132176 134810 contig of 2635 bp in length  
134911 136733 contig of 1823 bp in length  
136834 138736 contig of 1903 bp in length  
138837 141319 contig of 2483 bp in length  
141420 143860 contig of 2441 bp in length  
143961 145965 contig of 2005 bp in length  
145966 147716 contig of 1651 bp in length  
147817 149568 contig of 1752 bp in length  
149669 151034 contig of 1366 bp in length  
151135 153329 contig of 2195 bp in length  
153430 155008 contig of 1579 bp in length  
155109 156283 contig of 1175 bp in length  
156384 157915 contig of 1532 bp in length  
158016 161187 contig of 1698 bp in length  
161288 162633 contig of 1374 bp in length  
162734 163954 contig of 1221 bp in length  
163955 165397 contig of 1343 bp in length  
165498 166501 contig of 1004 bp in length  
166602 167634 contig of 1033 bp in length  
167735 168735 contig of 1001 bp in length

Sequence updated (26-May-2000).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 42 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12685: contig of 12685 bp in length  
12686 12786: gap of 100 bp  
12787 21378: contig of 8593 bp in length  
21379 31812: gap of 100 bp  
31813 40293: contig of 8381 bp in length  
40294 50693: gap of 100 bp  
50694 57517: contig of 6724 bp in length  
57518 65718: gap of 100 bp  
65719 72757: contig of 6939 bp in length  
72758 80449: gap of 100 bp  
80450 87968: contig of 7419 bp in length  
87969 93179: contig of 5111 bp in length  
93180 99470: contig of 6191 bp in length  
99471 105119: gap of 100 bp  
105120 109327: contig of 4108 bp in length  
109328 114708: gap of 100 bp  
114709 118902: contig of 4094 bp in length  
118903 123406: contig of 4404 bp in length  
123407 126415: gap of 100 bp  
126416 129075: gap of 100 bp  
129076 130072: contig of 897 bp in length  
130073 132075: gap of 100 bp  
132076 134810: contig of 2635 bp in length  
134811 136733: gap of 100 bp  
136734 138736: contig of 1903 bp in length  
138737 141319: contig of 2483 bp in length  
141320 143860: contig of 2441 bp in length  
143861 145965: gap of 100 bp  
145966 147716: contig of 1651 bp in length  
147717 149568: contig of 1752 bp in length  
149569 151034: contig of 1366 bp in length  
151035 153329: contig of 2195 bp in length  
153330 155008: contig of 1579 bp in length  
155009 156283: contig of 1175 bp in length  
156284 157915: contig of 1532 bp in length  
157916 161187: gap of 100 bp  
161188 162633: contig of 1374 bp in length  
162634 163954: contig of 1221 bp in length  
163955 165397: contig of 1343 bp in length  
165398 166501: contig of 1004 bp in length  
166502 167634: contig of 1033 bp in length

```

* 167635 167734: gap of 100 bp
* 167735 168735: contig of 1001 bp in length.
FEATURES
    source
        1..168735
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="18"
            /map="18q23"
            /clone="RP11-849119"
            /c=.12585
            /note="assembly_fragment"
        misc_feature
            12786..21378
            /note="assembly_fragment"
        misc_feature
            21479..31812
            /note="assembly_fragment"

Query Match      55.9%; Score 21.8; DB 14; Length 168735;
Best Local Similarity 78.8%; Pred. NO. 1.1e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATAGGAGCGTCTAATCTAGACACTCAGC 33
Db 160399 ATCGCCACCATCTAATCTAGACCTTCTCAG 160431

RESULT 14
AC021857/c
LOCUS AC021857 177575 bp DNA linear HTG 27-MAR-2001
DEFINITION Homo sapiens chromosome RPCI-11 clone RP11-563B11, WORKING DRAFT
SEQUENCE, 12 unordered pieces.
ACCESSION AC021857
VERSION AC021857.4 GI:9966354
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominiidae; Homo.
REFERENCE 1 (bases 1 to 177575)
    Waterston,R.H.
    The sequence of Homo sapiens clone
    Unpublished
REFERENCE 2 (bases 1 to 177575)
    Waterston,R.H.
    Direct Submission
    Submitted (20-JAN-2000) Genome Sequencing Center, Washington
    University School of Medicine, 4444 Forest Park Parkway, St. Louis,
    MO 63108, USA
COMMENT On Sep 4, 2000 this sequence version replaced gi:8887041.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0563B11
----- Summary Statistics -----
Sequencing vector: M13; 87%
Chemistry: Dye-terminator; plasmid; 13%
Chemistry: Dye-terminator Big Dye; 13% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168674 bases at least Q40
Consensus quality: 171499 bases at least Q30
Consensus quality: 173203 bases at least Q20
Insert size: 154000; agarose-gel
Insert size: 176475; sum-of-contigs
Quality coverage: 7.97 in Q20 bases; agarose-gel
Quality coverage: 7.62 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces

```

---

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1395: contig of 1395 bp in length
* 1396: gap of unknown length
* 1496: contig of 1306 bp in length
* 2801: gap of unknown length
* 2901: gap of unknown length
* 2902: contig of 1424 bp in length
* 4325: gap of unknown length
* 4425: contig of 1264 bp in length
* 5689: gap of unknown length
* 5789: contig of 1903 bp in length
* 7692: gap of unknown length
* 7792: contig of 2619 bp in length
* 10411: gap of unknown length
* 10511: contig of 1853 bp in length
* 12364: gap of unknown length
* 12464: contig of 2762 bp in length
* 15226: gap of unknown length
* 15326: contig of 11934 bp in length
* 17260: gap of unknown length
* 17360: contig of 17606 bp in length
* 27261: gap of unknown length
* 44967: contig of 19194 bp in length
* 45066: gap of unknown length
* 64260: contig of 19194 bp in length
* 64261: gap of unknown length
* 64360: contig of 113215 bp in length.
* 177575: Location/Qualifiers
FEATURES
    source
        1..177575
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="RP11-563B11"
            /clone="RP11-563B11"
            /c=.1395
            /note="assembly_name:Contig18"
        gap
            1396..1495
            /estimated_length=unknown
        misc_feature
            1496..2801
            /note="assembly_name:Contig25"
        gap
            2802..2901
            /estimated_length=unknown
        misc_feature
            2902..4325
            /note="assembly_name:Contig27"
        gap
            4326..4425
            /estimated_length=unknown
        misc_feature
            4426..5689
            /note="assembly_name:Contig28"
        gap
            5690..5789
            /estimated_length=unknown
        misc_feature
            5790..7692
            /note="assembly_name:Contig29"
        gap
            7693..7792
            /estimated_length=unknown
        misc_feature
            7793..10411
            /note="assembly_name:Contig30"
        gap
            10412..10511
            /estimated_length=unknown
        misc_feature
            10512..12364
            /note="assembly_name:Contig31"
        gap
            12365..12464
            /estimated_length=unknown
        misc_feature
            12465..15226
            /note="assembly_name:Contig32"
        gap
            15227..15326
            /estimated_length=unknown
        misc_feature
            15327..27260
            /note="assembly_name:Contig33"
        gap
            27261..27360
            /estimated_length=unknown
        misc_feature
            27361..44966
            /estimated_length=unknown

```

```

/note="assembly_name:Contig34
clone_end:17
vector_side:left"
gap
/estimated length=unknown
44967. .45066
/misc_feature
5067. .64360
/note="assembly_name:Contig35
clone_end:SP6
vector_side:left"
gap
54261. .64360
/estimated length=unknown
64361. .177575
/note="assembly_name:Contig36"

ORIGIN
Query Match 55.9%; Score 21.8; DB 14; Length 177575;
Best Local Similarity 78.8%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATAGCGAGGCTTAATTCAGAACACTCAGC 33
Db 124496 ATCGCCACCATCAATTCAGAACCTTCTCAGC 124464

RESULT 15
AC099689/c
LOCUS AC099689 201519 bp DNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens chromosome 18, clone RP11-849119, complete sequence.
ACCESSION AC099689
VERSION AC099689.4 GI:22655826
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 201519)
Birren,B., Nusbaum,C. and Lander,B.
Homo sapiens chromosome 18, clone RP11-849119
Unpublished
2 (bases 1 to 201519)
Birren,B., Linton,L., Nusbaum,C., Lander,B., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Riese,C., Rogov,P.,
Raymond,C., Retta,R., Riley,R., Rouse,C., Santos,R., Schauer,S.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (18-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 201519)
Birren,B., Nusbaum,C., Lander,B., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Riese,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 3, 2002 this sequence version replaced gi:22532588.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L13214
Center clone name: 849_I_19

----- Location/Qualifiers
1. 201519
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-849119"
/clone_lib="RPC1-11 Human Male BAC"
1. 2255
/rpt_family="L1"
repeat_region complement(248..379)
/rpt_family="L1PA15-16"
373..1421
/rpt_family="L1"
repeat_region 1422..1915
/rpt_family="L1PA16"
1921..2512
/rpt_family="L1MBC"
2653..2928
/rpt_family="L1MBC"
repeat_region complement(2929..3505)
/rpt_family="L1MBC1"

```

```
repeat_region 3506..3708
/rpt_family="L1MEc"
repeat_region 3840..4264
/rpt_family="MLT2P"
repeat_region 4271..4502
/rpt_family="L1ME3"
repeat_region 5703..6155
/rpt_family="L1MB7"
repeat_region 8041..8164
/rpt_family="MIR"
repeat_region 10799..10967
/rpt_family="(ATGGTG)n"
repeat_region 11369..11393
/rpt_family="(ATG)n"
repeat_region complement(11798..11892)
/rpt_family="MLT1J2"
repeat_region complement(13779..13888)
/rpt_family="L2"
repeat_region 17923..18222
/rpt_family="AluJb"
repeat_region 19455..19760
/rpt_family="AluJo"
repeat_region 20132..20302
/rpt_family="AluYa5"
repeat_region 20303..20488
/rpt_family="AluSg/x"
repeat_region 20498..20786
/rpt_family="AluX"
repeat_region 20810..20980
/rpt_family="FRAM"
repeat_region complement(23080..23364)
/rpt_family="AluJb"
repeat_region 23903..24105
/rpt_family="(CCCCAG)n"
repeat_region 24152..24211
/rpt_family="(CCCCAG)n"
repeat_region 24244..24433
/rpt_family="(CCCCAG)n"
repeat_region 24475..24653
/rpt_family="(CCCCAG)n"
repeat_region 25987..26093
/rpt_family="Charlie4"
repeat_region 26229..26331
/rpt_family="Charlie4"
repeat_region 28009..28083
/rpt_family="MIR"
repeat_region complement(28483..28723)
/rpt_family="MIR3"
repeat_region 28821..28901
/rpt_family="(TAGA)n"
repeat_region 28906..29107
/rpt_family="(TGGA)n"
repeat_region 29122..29323
/rpt_family="(TGGA)n"
repeat_region 29338..29539
/rpt_family="(TGGA)n"
repeat_region 29407..29560
/rpt_family="(TGGA)n"
unsure /note="single clone coverage"
unsure 29407..29423
repeat_region /note="< 30 qual single clone coverage"
repeat_region 29554..30046
/rpt_family="(TGGA)n"
repeat_region 30016..30191
/rpt_family="(TAGA)n"
repeat_region 30194..30493
/rpt_family="AluYa5"
repeat_region 30494..30816
/rpt_family="(TAGA)n"
repeat_region complement(31718..31919)
/rpt_family="AluJ"
repeat_region complement(31933..32005)
/rpt_family="AluJ"
repeat_region 33370..33548
```

```
repeat_region /rpt_family="(CA)n"
33561..33649
/rpt_family="(CA)n"
repeat_region 33720..33827
/rpt_family="L2"
repeat_region 34885..35203
/rpt_family="L2"

Query Match 55.9%; Score 21.8; DB 8; Length 201519;
Best Local Similarity 78.8%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ATAGCGAGCGTCTAATTCCTAGACACTCACACG 33
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38836 ATCGCCACCATCTAATTCCTAGACCTTCTCAG 38804

Search completed: January 30, 2006, 02:34:22
Job time : 898 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2006, 22:23:56 ; Search time 219 Seconds  
(without alignments)  
1186.863 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3

Perfect score: 39  
Sequence: 1 atagcgcgcgtcattctagacactcacacgcgtctt 39

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_21:\*

1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*  
14: Geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.4	57.4	653	2	AAV52043 Helicobac
2	20.6	52.8	2409	13	AdS47793 Bacterial
3	20.2	51.8	402	4	Aal19396 Human bre
4	20.2	51.8	402	4	Aal09009 Human bre
5	20.2	51.8	451	11	Acn79225 Breast ca
6	20.2	51.8	523	4	Aal09684 Human bre
7	20.2	51.8	561	4	Aak92531 Human cDN
8	20.2	51.8	561	12	AdL28958 3' end of
9	20.2	51.8	603	4	Aal18356 Human bre
10	20.2	51.8	686	11	Acn79884 Breast ca
11	20.2	51.8	1142	4	Aas33195 DNA encod
12	20.2	51.8	2709	4	Aak94205 Human ful
13	20.2	51.8	2709	12	AdL30735 Full leng
14	20.2	51.8	2759	4	Aak53167 Human pol
15	20.2	51.8	2759	4	Aal59801 Human pol
16	20.2	51.8	2844	4	Aak52183 Human pol
17	20.2	51.8	2844	4	Aal58015 Human pol
18	20.2	51.8	2864	12	AdQ86453 Human tum
19	20.2	51.8	2871	13	Adr99114 GALNT6 co

c	20	20.2	51.8	2881	6	ABT07725	Abt07725 Breast ca
c	21	20.2	51.8	2881	11	ACN44271	Acn44271 Human mRN
c	22	20.2	51.8	2881	12	ADN04939	Adn04939 Antipsori
c	23	20.2	51.8	2881	14	AEb22801	Aeb22801 Human col
c	24	20.2	51.8	3230	11	ACN89199	Acn89199 Breast ca
c	25	20.2	51.8	12216	4	AAS33427	Aas33427 DNA encod
c	26	20.2	51.8	46404	11	ACN44270	Acn44270 Human gen
c	27	20	51.3	20	3	AAO9497	Aao9497 Primer GW
c	28	20	51.3	934	2	AAx14231	Aax14231 H. pylori
c	29	20	51.3	1119	3	AAA09495	Aaa09495 H. pylori
c	30	20	51.3	110000	14	ADZ59507_2	Continuation (3 of
c	31	19.8	50.8	386	3	AAC01025	Aac01025 Human sec
c	32	19.8	50.8	569	5	ABV35629	Abv35629 Human pro
c	33	19.8	50.8	569	5	ABV44435	Abv44435 Human pro
c	34	19.8	50.8	570	5	ABV14550	Abv14550 Human pro
c	35	19.8	50.8	622	10	ADB57315	Adb57315 Toxicity-
c	36	19.8	50.8	622	10	ADB51871	Adb51871 Primary r
c	37	19.8	50.8	622	13	ADV40531	Adv40531 Rat cardi
c	38	19.8	50.8	697	2	Aaz33534	Aaz33534 Human pro
c	39	19.8	50.8	731	6	AAS62285	Aas62285 cDNA sequ
c	40	19.8	50.8	790	6	AAL41606	Aal41606 Human col
c	41	19.8	50.8	800	3	AAC99839	Aac99839 Human sec
c	42	19.8	50.8	954	6	ABQ54895	Abq54895 Human ova
c	43	19.8	50.8	1046	13	ADR25781	Adr25781 Breast ca
c	44	19.8	50.8	1062	13	ADS09846	Ads09846 Human the
c	45	19.8	50.8	1249	5	ABV23395	Abv23395 Human pro

## ALIGNMENTS

## RESULT 1

AAV52043	ID	AAV52043 standard; DNA; 653 BP.
XX	AC	AAV52043;
XX	DT	09-NOV-1998 (first entry)
XX	DE	Helicobacter polypeptide GHPO 1273 DNA.
XX	KW	GHPO 1273; infection; therapy; diagnosis; vaccine; gastritis; ulcer; ss.
XX	OS	Helicobacter pylori.
XX	PH	Key
XX	FT	CDS
XX	XX	Location/Qualifiers
XX	XX	63..593
XX	XX	/*tag= a
XX	PN	WO9821225-A1.
XX	PD	22-MAY-1998.
XX	PF	14-NOV-1997; 97WO-US021353.
XX	PR	14-NOV-1996; 96US-00749051.
XX	PR	01-APR-1997; 97US-00831309.
XX	PR	01-APR-1997; 97US-00833457.
XX	PR	01-APR-1997; 97US-00834705.
XX	PR	24-JUN-1997; 97US-00881227.
XX	PR	29-JUL-1997; 97US-00902615.
XX	PA	(INNR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX	PA	(PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX	PI	(HUMA-) HUMAN GENOME SCI INC.
XX	PI	Haas R, Kleanthous H, Tomb J, Miller C, Al-Garawi A, Odenbreit S;
XX	PI	Meyer T;
XX	DR	WPI; 1998-297855/26.
XX	DR	P-PSDB; AAW71508.
XX	PT	Helicobacter polynucleotide and polypeptide sequences - useful to treat

or prevent gastrointestinal infection.

XX PS Disclosure; Page 174-175; 362pp; English.

XX CC This polynucleotide encodes Helicobacter pylori polypeptide GHPO 1273 (see AAW71508). 85 Polynucleotides (see AAV52009-V52093) encoding claimed Helicobacter polypeptides (see AAW71474-W71558) have been identified in the H. pylori genome using a transposon shuttle mutagenesis method in which TnMax9 mini-blaM was used for insertional mutagenesis of an H. pylori gene library established in Escherichia coli. The invention provides methods for producing these polypeptides in recombinant host systems, and related expression cassettes, vectors and transformed or transfected host cells; live vaccine vectors that contain polynucleotides of the invention and which can be used to prevent or treat Helicobacter infection; therapeutic and/or prophylactic methods involving administration of polynucleotide molecules (either in naked form or delivered with a delivery vehicle), polypeptides or monospecific antibodies; methods for detecting the presence of Helicobacter in samples using e.g. polynucleotide molecules; and methods for purifying polypeptides of the invention

XX SQ Sequence 653 BP; 234 A; 126 C; 104 G; 189 T; 0 U; 0 Other;

Query Match 57.4%; Score 22.4; DB 2; Length 653;  
Best Local Similarity 95.8%; Pred. No. 11;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 16 TTCTAGACACTCAGCGGCTTT 39  
Db 547 TGCTAGAACACTCAGCGGCTTT 570

RESULT 2  
ADS47793/C  
ID ADS47793 standard; cDNA; 2409 BP.  
XX AC ADS47793;  
XX DT 02-DEC-2004 (first entry)  
XX DE Bacterial polynucleotide #2536.  
XX KW Recombinant DNA construct; transformed plant; improved plant property; cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis; pathogen tolerance; pest tolerance; plant disease resistance; cell cycle pathway modification; plant growth regulator; homologous recombination; seed oil yield; protein yield; carbohydrate; nitrogen; phosphorus; photosynthesis; lignin; galactomannan; bacterial polynucleotide; gene; ss.  
XX OS Bacteria.  
XX PN US2002333675-A1.  
XX PD 18-DEC-2003.  
XX PF 20-FEB-2003; 2003US-00369493.  
XX PR 21-FEB-2002; 2002US-0360039P.  
XX PA (CAOY/) CAO Y.  
XX PA (HINK/) HINKLE G J.  
XX PA (SLATY/) SLATER S C.  
XX PA (CHEN/) CHEN X.  
XX PA (GOLD/) GOLDMAN B S.  
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX FT New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 26223; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polynucleotide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 2409 BP; 741 A; 461 C; 555 G; 652 T; 0 U; 0 Other;

Query Match 52.8%; Score 20.6; DB 13; Length 2409;  
Best Local Similarity 74.3%; Pred. No. 88;  
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 GCGAGCGCTTAATCTTAGACACTCAGCGGCTT 38  
Db 534 GAGAGCGCTTAATCTTAGACACTCAGCGGCTT 500

RESULT 3  
AAL19396  
ID AAL19396 standard; cDNA; 402 BP.  
XX AC AAL19396;  
XX DT 07-DEC-2001 (first entry)  
XX DE Human breast cancer expressed polynucleotide 11853.  
XX KW Human; breast cancer; cell marker; cytostatic; ss.  
XX OS Homo sapiens.  
XX PN WO200151628-A2.  
XX PD 19-JUL-2001.  
XX PF 10-JAN-2001; 2001WO-US0000798.  
XX PR 14-JAN-2000; 2000US-0176077P.  
XX PR 14-MAR-2000; 2000US-0189167P.  
XX PR 24-MAR-2000; 2000US-0192099P.  
XX PR 29-MAR-2000; 2000US-0193480P.  
XX PR 15-MAY-2000; 2000US-0205230P.  
XX PR 09-JUN-2000; 2000US-0211315P.  
XX PR 25-JUL-2000; 2000US-0220534P.  
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;  
XX WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 2106; 3695pp; English.

PS The invention relates to human breast cancer expressed polynucleotides (AA07544-AA026789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

XX Sequence 402 BP; 88 A; 99 C; 110 G; 105 T; 0 U; 0 Other;

XX Query Match 51.8%; Score 20.2; DB 4; Length 402;

XX Best Local Similarity 75.8%; Pred. No. 96;

XX Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTTAGAACACTCACACGC 34

Db 62 TTGGAGCGCTCTATTCTTAGAACAGGAAACGC 94

RESULT 4

AA09009

ID AAL09009 standard; cDNA; 402 BP.

XX AC AAL09009;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 1466.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 307; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides (AA07544-AA026789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

SQ Sequence 402 BP; 85 A; 100 C; 114 G; 103 T; 0 U; 0 Other;

XX Query Match 51.8%; Score 20.2; DB 4; Length 402;

XX Best Local Similarity 75.8%; Pred. No. 96;

XX Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTTAGAACACTCACACGC 34

Db 77 TTGGAGCGCTCTATTCTTAGAACAGGAAACGC 109

RESULT 5

ACN79225

ID ACN79225 standard; DNA; 451 BP.

XX AC ACN79225;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 375.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for detecting presence of polypeptide in sample, as a marker for breast cancer.

XX Disclosure; SEQ ID NO 375; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with breast cancer which is encoded by a nucleic acid molecule comprising a nucleotide sequence (S1). Further disclosed is an antibody that binds to the polypeptide of the invention. The activity of the polypeptide of the invention may be described as cytostatic. The antibody is useful for detecting the presence of (I) in a sample. Nucleic acid molecules of the invention are useful in the detection of breast tumours. (I) is useful as a marker for breast cancer and in breast cancer therapy. Sequences given in records ACN78851-ACN92934 represent nucleic acid markers associated with breast cancer. Note: The sequence listing does not form part of the specification but may be obtained in electronic format from the USPTO web site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 451 BP; 98 A; 112 C; 127 G; 113 T; 0 U; 1 Other;

XX Query Match 51.8%; Score 20.2; DB 11; Length 451;

XX Best Local Similarity 75.8%; Pred. No. 98;

XX Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTTAGAACACTCACACGC 34

Db 109 TTGGAGCGCTCTATTCTTAGAACAGGAAACGC 141

RESULT 6

AA09684

ID AAL09684 standard; cDNA; 523 BP.

XX

```
AC AAL09694;
XX
XX
DT 07-DEC-2001 (first entry)
DE Human breast cancer expressed polynucleotide 2141.
XX
XX Human; breast cancer; cell marker; cytotstatic; ss.
XX
XX Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX
XX 14-MAR-2000; 2000US-0189167P.
XX
XX 24-MAR-2000; 2000US-0192099P.
XX
XX 29-MAR-2000; 2000US-0193480P.
XX
XX 15-MAY-2000; 2000US-0205230P.
XX
XX 09-JUN-2000; 2000US-0211315P.
XX
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 413; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AAL07544-AAL26789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterising treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity
XX
XX SQ Sequence 523 BP; 110 A; 120 C; 151 G; 122 T; 0 U; 20 Other;

Query Match 51.8%; Score 20.2; DB 4; Length 523;
Best Local Similarity 75.8%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTAGAACACTCACAGC 34
Db 98 TTGGAGCGCTCTATTCTTAGAACAGGAAACGC 130

RESULT 7
AAK92531
ID AAK92531 standard; cDNA; 561 BP.
XX
XX AAK92531;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human cDNA 3'-end sequence, SEQ ID NO: 991.
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX Homo sapiens.
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX

PF 07-JUL-2000; 2000EP-00114089.
XX
XX 08-JUL-1999; 99JP-00194486.
XX
XX 11-JAN-2000; 2000JP-00118774.
XX
XX 02-MAY-2000; 2000JP-00183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
XX WPI; 2001-524255/58.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
XX in genetic manipulation.
XX
XX Claim 3; SEQ ID NO 991; 1380pp + Sequence Listing; English.
XX
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesising the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is the nucleotide sequence of the 3'-end of
XX a cDNA provided in the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in CD-
XX ROM format directly from EPO
XX
XX SQ Sequence 561 BP; 132 A; 146 C; 133 G; 146 T; 0 U; 4 Other;

Query Match 51.8%; Score 20.2; DB 4; Length 561;
Best Local Similarity 75.8%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTCTAGAACACTCACAGC 34
Db 369 TTGGAGCGCTCTATTCTTAGAACAGGAAACGC 401

RESULT 8
ADL28958
ID ADL28958 standard; cDNA; 561 BP.
XX
XX ADL28958;
XX
XX 20-MAY-2004 (first entry)
XX
XX 3' end of a human cDNA molecule SeqID 991.
XX
XX human; medicine; signal transduction; glycoprotein; transcription;
XX oligo-capping method; ss.
XX
XX Homo sapiens.
XX
XX EP1396543-A2.
XX
XX 10-MAR-2004.
XX
XX 07-JUL-2000; 2003EP-00025638.
XX
XX 08-JUL-1999; 99JP-00194486.
XX
XX 11-JAN-2000; 2000JP-00118774.
XX
XX 02-MAY-2000; 2000JP-00183865.
XX
XX 07-JUL-2000; 2000EP-00114089.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
```



DR WPI; 2004-204755/20.  
 XX  
 PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full  
 PT length human cDNAs.  
 XX  
 PS Disclosure; SEQ ID NO 991; 1340pp; English.  
 XX  
 CC This invention relates to a novel primers useful for synthesizing full  
 CC length cDNA molecules that encode human proteins. Specifically, it refers  
 CC to secretory or membrane proteins that are potential therapeutic agents/  
 CC target molecules in the field of medicine, and in particular genes  
 CC encoding proteins that are associated with signal transduction.  
 CC glycoproteins and transcription. The present invention describes a method  
 CC for efficiently cloning a full length human cDNA from both the 5' and 3',  
 CC ends using the oligo-capping method. This polynucleotide sequence is the  
 CC 3' end of a full length human cDNA sequence of the invention.  
 XX  
 SQ Sequence 561 BP; 132 A; 146 C; 133 G; 146 T; 0 U; 4 Other;

Query Match 51.8%; Score 20.2; DB 12; Length 561;  
 Best Local Similarity 75.8%; Pred. No. 1e+02;  
 Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 TAGGAGCGCTCAATTCCTAGACACTCACACGC 34  
 | | | | | | | | | | | | | | | | | | | | | |  
 Db 369 TTGGGAGCGCTCATTTCTAGAACAGGAAACGC 401

RESULT 9  
 AAL18356  
 ID AAL18356 standard; cDNA; 603 BP.  
 XX  
 AC AAL18356;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 DE Human breast cancer expressed polynucleotide 10813.  
 XX  
 KW Human; breast cancer; cell marker; cytostatic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200151628-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 10-JAN-2001; 2001WO-US000798.  
 XX  
 PR 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0192099P.  
 PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.  
 XX  
 (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA  
 PI Lillie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 DR WPI; 2001-451856/48.  
 XX  
 PT New peptide useful as a marker for the diagnosis of breast cancer.  
 XX  
 PS Claim 1; Page 1925; 3695pp; English.  
 XX  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterizing treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded

AAS33195/c  
ID AAS33195 standard; cDNA; 1142 BP.  
XX  
AC AAS33195;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE DNA encoding human secreted protein, Seq ID No 154.  
XX  
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytoskeletal; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155326-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001347.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216664P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0236279P.  
PR 22-AUG-2000; 2000US-0236681P.  
PR 22-AUG-2000; 2000US-0236688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241807P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.



```
XX 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183865.
PR 07-JUL-2000; 2000EP-00114089.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PA
XX Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX P-PSDB; ADL30736.
DR WPI; 2004-204755/20.
XX PT
XX New oligonucleotide primers (830 cDNAs) useful for synthesizing full
PT length human cDNAs.
XX Example 1; SEQ ID NO 2768; 1340pp; English.
XX This invention relates to a novel primers useful for synthesising full
CC length cDNA molecules that encode human proteins. Specifically, it refers
CC to secretory or membrane proteins that are potential therapeutic agents/
CC target molecules in the field of medicine, and in particular genes
CC encoding proteins that are associated with signal transduction,
CC glycoproteins and transcription. The present invention describes a method
CC for efficiently cloning a full length human cDNA from both the 5' and 3'
CC ends using the oligo-capping method. This polynucleotide sequence is a
CC full length human cDNA clone of the invention.
XX SQ
SQ Sequence 2709 BP; 658 A; 756 C; 760 G; 535 T; 0 U; 0 Other;
Query Match 51.8%; Score 20.2; DB 12; Length 2709;
Best Local Similarity 75.8%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TAGCGAGCGCTCTAATTTCTAGAACACTCTCACGCG 34
DB 2341 TTGGGAGCGCTCTAATTTCTAGAACAGGAAACGCG 2309
RESULT 14
AAK53167/c
ID AAK53167 standard; cDNA; 2759 BP.
XX AC
XX AAK53167;
XX DT
XX 06-NOV-2001 (first entry)
XX DE
XX Human polynucleotide SEQ ID NO 2696.
XX KW
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX OS
XX Homo sapiens.
XX FN
XX WO200157190-A2.
XX PD
XX 09-AUG-2001.
XX XX
XX 05-FEB-2001; 2001WO-US004098.
XX PR
XX 03-FEB-2000; 2000US-00496914.
XX PR 27-APR-2000; 2000US-00560875.
XX PR 20-JUN-2000; 2000US-00598075.
XX PR 19-JUL-2000; 2000US-00620325.
XX PR 01-SEP-2000; 2000US-00654936.
XX PR 15-SEP-2000; 2000US-00663561.
XX PR 20-OCT-2000; 2000US-00693325.
XX PR 30-NOV-2000; 2000US-00728422.
XX PA
(HYSE-) HYSEQ INC.
```

```
XX Tang YT, Liu C, Dzmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX P-PSDB; AAM80034.
XX PT
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX Claim 1; Page 4933-4934; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX SQ
SQ Sequence 2759 BP; 662 A; 771 C; 781 G; 545 T; 0 U; 0 Other;
Query Match 51.8%; Score 20.2; DB 4; Length 2759;
Best Local Similarity 75.8%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TAGCGAGCGCTCTAATTTCTAGAACACTCTCACGCG 34
DB 2400 TTGGGAGCGCTCTAATTTCTAGAACAGGAAACGCG 2368
RESULT 15
AAI59801/c
ID AAI59801 standard; cDNA; 2759 BP.
XX AC
XX AAI59801;
XX DT
XX 22-OCT-2001 (first entry)
XX DE
XX Human polynucleotide SEQ ID NO 3790.
XX KW
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX OS
XX Homo sapiens.
XX FN
XX WO200153312-A1.
XX PD
XX 26-JUL-2001.
XX XX
XX 26-DEC-2000; 2000WO-US034263.
XX PR
XX 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX XX
```

PA (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR P-PSDB; AAM40845.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
PS Claim 1; SEQ ID NO 3790; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX  
SQ Sequence 2759 BP; 662 A; 771 C; 781 G; 545 T; 0 U; 0 Other;

Query Match 51.8%; Score 20.2; DB 4; Length 2759;  
Best Local Similarity 75.8%; Pred. No. 1.3e+02;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 TAGCGAGCGTCTAATTTCTAGAACACTCACAACGC 34  
DB 2400 TTGGGAGCGCTCTAATTTCTAGAACAGGAAACGC 2368

Search completed: January 30, 2006, 02:04:26  
Job time : 222 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6

Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:24:06 ; Search time 1556 Seconds  
(without alignments)  
1172.684 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3

Perfect score: 39

Sequence: 1 atagcgagcgtctaattctagaaactcacacgcgtctt 39

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 0.1

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_hic:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_est7:\*

9: gb\_gss1:\*

10: gb\_gss2:\*

11: gb\_gss3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24.2	62.1	560	10 CNS011EX	AL145642 Anopheles
2	22.6	57.9	1264	10 CL473630	CL473630 SAIL_204
3	22.6	57.9	1277	10 CL495547	CL495547 SAIL_670
c 4	22.4	57.4	831	9 AQ871927	AQ871927 nbe0045B
5	22.2	56.9	264	11 CNS04LUL	AL296598 Tetraodon
6	22.2	56.9	850	10 AG136295	AG136295 Pah trogl
c 7	22.2	56.4	506	1 A1179396	A1179396 EST223098
8	21.8	55.9	439	9 AQ494887	AQ494887 HS_5195_A
9	21.8	55.9	635	9 CE120785	CE120785 tigr-gss-
c 10	21.8	55.9	935	7 CO731803	CO731803 JBM027A05
11	21.8	55.9	1032	10 CL491525	CL491525 SAIL_557
12	21.4	54.9	295	1 AM407061	AM407061 UI-HF-B10
13	21.4	54.9	325	3 BM851087	BM851087 K-EST0131
14	21.4	54.9	445	3 BM851189	BM851189 K-EST0131
c 15	21.4	54.9	504	1 AA695264	AA695264 GM02505.5
c 16	21.4	54.9	565	5 BU580658	BU580658 in36e09.x
17	21.4	54.9	807	10 CW975483	CW975483 A1AA-aaa8
18	21.4	54.9	863	6 CF660787	CF660787 CCLM09a32
19	21.2	54.4	479	2 BE162200	BE162200 IL2-HT044
20	21.2	54.4	598	3 BJ192483	BJ192483 BJ192483
c 21	21.2	54.4	756	8 CX104655	CX104655 B1103L08
22	21	53.8	437	2 BF808552	BF808552 CM1-CI009

c 23	21	53.8	588	8	CV952286
24	21	53.8	590	10	CE532527
c 25	21	53.8	895	7	CNO17674
26	21	53.8	953	10	CL516085
27	21	53.8	1138	2	BG852398
28	20.8	53.3	571	10	CNS01ZYZ
c 29	20.8	53.3	668	10	CE629311
30	20.8	53.3	699	9	BH965866
c 31	20.8	53.3	744	7	CV673875
32	20.8	53.3	840	9	BH661798
c 33	20.8	53.3	841	2	BG367933
c 34	20.6	52.8	476	9	AO685782
35	20.6	52.8	504	7	CE537994
36	20.6	52.8	665	2	BB410213
37	20.6	52.8	775	8	CV991159
c 38	20.6	52.8	841	10	CZ380115
39	20.6	52.8	993	2	BG334170
40	20.6	52.8	1011	10	DU015562
c 41	20.6	52.8	1309	5	BU367266
42	20.6	52.8	1469	10	AG065308
c 43	20.4	52.3	234	10	CW069414
44	20.4	52.3	295	10	CL262429
c 45	20.4	52.3	315	2	BB531263

## ALIGNMENTS

RESULT 1	CNS011EX	560 bp	DNA	linear	GSS 12-JUN-2001
LOCUS	Anopheles gambiae GSS SP6 end of clone 10005 of NotreDame1 library				
DEFINITION	from strain PEST of Anopheles gambiae (African malaria mosquito), genomic survey sequence.				
ACCESSION	AL145642				
VERSION	AL145642.1				
KEYWORDS	GSS.				
SOURCE	Anopheles gambiae (African malaria mosquito)				
ORGANISM	Anopheles gambiae				
REFERENCE	1 (bases 1 to 560)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
REFERENCE	2 (bases 1 to 560)				
AUTHORS	Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France				
COMMENT	This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.				
FEATURES	Location/Qualifiers				
source	1. 560				
	/organism="Anopheles gambiae"				
	/mol_type="genomic DNA"				
	/strain="PEST"				
	/db_xref="taxon:7165"				
	/clone="10005"				
	/clone.lib="NotreDame1"				
	/note="end : SP6"				
ORIGIN					
Query Match	62.1%	Score	24.2	DB	10
Best Local Similarity	78.4%	Pred. No.	15		
Matches	29	Conservative	0	Mismatches	8
		Indels	0	Gaps	0
QY	3	ACGAGAGCGTCTAATTCTAGAACACTCACACGCGTCTT	39		







Qy	3	AGCGAGCGTCTAATTTAGAACACTCACAGCGTTC	37
Db	627	ACCAGCGCTCAATATACCAAAACACTCACAGCGCTC	661
RESULT 7			
LOCUS	AI179396/c		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST 08-JAN-1999
ACCESSION	RPSCG66.3'	end, mRNA sequence.	
VERSION	AI179396		
KEYWORDS	AI179396.1	GI:4135155	
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 506)		
TITLE	Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J., Kervavage,A.R. and Adams,M.D.		
JOURNAL	Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat		
COMMENT	Gene Index		
FEATURES	source		
LOCUS	AI179396		
DEFINITION	EST223098 Normalized rat spleen, Bento Soares Rattus sp. cDNA clone	506 bp	linear mRNA EST

```

/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 55.9%; Score 21.8; DB 9; Length 635;
Best Local Similarity 92.0%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 GTCTAATTCAGACACTCACACGC 34
Db 195 GTCCAAATTCAGACACTCACAC 219

RESULT 10
LOCUS CO731803/c
DEFINITION JBW027A05.b_047.abi Pineapple green mature fruit cDNA library
Ananas comosus cDNA clone JBW027A05 similar to 27k
vesicle-associated membrane protein-associated protein, mRNA
sequence.
ACCESSION CO731803
VERSION CO731803.1 GI:56848975
KEYWORDS EST.
SOURCE Ananas comosus (pineapple)
ORGANISM Ananas comosus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Bromeliaceae;
Ananas.
REFERENCE 1 (bases 1 to 935)
AUTHORS Moyle,R., Fairbairn,D.J., Ripi,J., Crowe,M. and Botella,J.R.
TITLE Developing pineapple fruit has a small transcriptome dominated by
metallothionein
JOURNAL J. Exp. Bot. 56 (409), 101-112 (2005)
PUBMED 15520025
COMMENT Contact: Crowe ML
School of Life Sciences
The University of Queensland
Brisbane, Queensland 4072, Australia
Tel: +617 33462621
Email: m.crowe@imb.uq.edu.au
Insert Length: 935 Std Error: 0.00.
Location/Qualifiers
1..935
/organism="Ananas comosus"
/mol_type="mRNA"
/cultivar="smooth cayenne"
/db_xref="taxon:4615"
/clone="JBW027A05"
/tissue_type="Fruit flesh"
/dev_stages="Mature green fruit"
/clone_lib="Pineapple green mature fruit cDNA library"

ORIGIN
Query Match 55.9%; Score 21.8; DB 7; Length 935;
Best Local Similarity 78.8%; Pred. No. 1.8e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAGCGAGCGTCTTAATTCAGACACTCACACGC 34
Db 897 TAGTGTGTTCTTAATTCAGACACTCACAC 865

RESULT 11
LOCUS CL491525
DEFINITION SAIL_557_D07.v1 SAIL Collection Arabidopsis thaliana genomic clone
SAIL_557_D07.v1, genomic survey sequence.
ACCESSION CL491525
VERSION CL491525.1 GI:45976434

```

```

KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1032)
AUTHORS Sessions,A., Burke,B., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
PUBMED 12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS823589; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..1032
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_557_D07.v1"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

ORIGIN
Query Match 55.9%; Score 21.8; DB 10; Length 1032;
Best Local Similarity 78.8%; Pred. No. 1.8e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 CGAGCGTCTAATTCAGACACTCACACGCCTC 37
Db 90 CGGCGCTCTAATTCAGTACATTAACAACGTC 122

RESULT 12
LOCUS AW407061
DEFINITION UI-HP-BL0-acy-b-01-0-UI.r1 NIH MGC_37 Homo sapiens cDNA clone
IMAGE:3060433 5', mRNA sequence.
ACCESSION AW407061
VERSION AW407061.1 GI:6926118
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 295)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```



**Qy**

1 ATAGCGAGCGTCTAAATTC TAGAACACTCACACGGGTCTT 39  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

**D6**

287 ATAGGGATCGTCC AACCCGAGAACATA CA AACTCCTCTT 325  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15	AA695264/C	LOCUS	AA695264	504 bp	mRNA	linear	EST 23-APR-2001
DEFINITION	GM02505.5prime GM Drosophila melanogaster ovary Bluescript Drosophila melanogaster cDNA clone GM02505 5, mRNA sequence.						

ACCESSION	AA695264	
KEYWORDS	AA695264.2	GI:13760555
VERSION	EST	

KEYWORDS	SOURCE	ORGANISM
ESSI.		
		Drosophila melanogaster (fruit fly)
		Drosophila melanogaster
		Drosophila melanogaster
		Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
		Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1. (bases 1 to 504)  
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,  
Lewis, S. and Rubin, G.M.  
BDGP/HMMI Drosophila EST Project  
Unpublished (2001)  
On Dec 18, 1997 this sequence version replaced gi:2697884.  
Other\_ESTs: GM02505.3prime  
Contact: Stapleton, M.

Lawrence Berkeley National Lab  
One Cyclotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
hit genomic AE003504; arm:X [16667865.16977278]  
estimated-cyto:15B4-15E7: 04/10/2001  
Plate: GM.25 row: A column: 5  
High quality sequence stop: 459

```

FEATURES
source
    Location/Qualifiers
      1..504
        /organism="Drosophila melanogaster"
        /mol_type="mRNA"
        /db_xref="taxon:7227"
        /clone="GW02505"
        /sex="female"
        /dev_stage="newly eclosed females: germanium-stage 6"
        /lab_host="SOLR"
        /clone_lib="GW Drosophila melanogaster ovary Bluescript"
        /note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
        Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
        Synthesis kit. Oligo df-primed and directionally cloned at
        EcoRI and XhoI in Bluescript SK(+/-)"

```

Query Match	54.9%;	Score 21.4;	DB 1;	Length 504;
Best Local Similarity	71.8%;	Pred. No. 2.5e+02;		
Matches 28;	Conservative	0;	Mismatches 11;	Indels 0;
			Gaps 0;	

**Qy**

1 ATAGCGGCGTCTAATTCTAGAACA CTACACGGGTCTT 39  
||||| ||||| || | ||||| ||| ||||| |||

**D6**

206 ATAACGAGCGTGATATATATA GAACA CATNACCACATATT 168  
||||| ||||| || | ||||| ||| ||||| |||

Search completed: January 30, 2006, 03:26:20  
Job time : 1560 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 01:49:51 ; Search time 73 Seconds  
(without alignments)  
949.656 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3

Perfect score: 39  
Sequence: 1 atacggcgctctattctagacactcacgcgcttt 39

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgm2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgm2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgm2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgm2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgm2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgm2\_6/ptodata/1/ina/PCUTUS COMB.seq.\*  
7: /cgm2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgm2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgm2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21.4	54.9	2182	3	US-09-270-767-13447
C 2	20.4	52.3	14641	3	US-09-949-016-15834
C 3	20	51.3	20	3	US-09-433-598-4
C 4	20	51.3	20	3	US-09-848-838A-4
5	20	51.3	1119	3	US-09-433-598-1
6	20	51.3	1119	3	US-09-848-838A-1
7	19.8	50.8	386	3	US-09-513-999C-1023
8	19.8	50.8	554	3	US-09-621-976-1260
9	19.6	50.3	1807	3	US-09-902-540-316
10	19	48.7	19	3	US-09-433-598-3
11	19	48.7	19	3	US-09-848-838A-3
C 12	19	48.7	4515	3	US-09-328-352-2320
C 13	19	48.7	8874	3	US-08-894-344C-1
14	19	48.7	8874	3	US-09-678-023A-1
15	19	48.7	9621	3	US-09-125-028-1
C 16	19	48.7	54161	3	US-09-943-016-11905
C 17	18.8	48.2	92581	3	US-09-949-016-12182
C 18	18.8	48.2	92581	3	US-09-949-016-16542
19	18.6	47.7	519	3	US-09-248-796A-2238
20	18.6	47.7	930	4	US-09-605-703B-2563
C 21	18.6	47.7	1215	3	US-09-248-796A-2972
22	18.6	47.7	2502	2	US-08-062-472B-2
C 23	18.6	47.7	17341	3	US-09-415-946-1
24	18.6	47.7	32042	3	US-09-245-281-44

Sequence 63, Appl  
Sequence 63, Appl  
Sequence 810, App  
Sequence 673, App  
Sequence 13210, A  
Sequence 37, Appl  
Sequence 37, Appl  
Sequence 264, App  
Sequence 15079, A  
Sequence 12763, A  
Sequence 16330, A  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 294, App  
Sequence 65208, A  
Sequence 83545, A  
Sequence 83546, A  
Sequence 83547, A  
Sequence 83548, A  
Sequence 6394, Ap

25 18.6 47.7 32042 3 US-09-340-620A-63  
26 18.6 47.7 32042 3 US-09-728-721-63  
27 18.6 47.7 39553 3 US-09-949-002-810  
28 18.6 47.7 58111 3 US-09-949-002-673  
29 18.6 47.7 109925 3 US-09-949-016-13210  
C 30 18.4 47.2 2190 2 US-09-036-582-37  
C 31 18.4 47.2 2190 3 US-09-318-141-37  
C 32 18.4 47.2 2819 3 US-10-104-047-264  
C 33 18.4 47.2 39690 3 US-09-949-016-15079  
C 34 18.4 47.2 52199 3 US-09-949-016-12763  
C 35 18.4 47.2 52203 3 US-09-949-016-16330  
C 36 18.4 47.2 70559 3 US-09-409-800B-1  
C 37 18.4 47.2 4403765 3 US-09-103-840A-2  
C 38 18.4 47.2 4411529 3 US-09-103-840A-1  
C 39 18.2 46.7 490 3 US-09-621-976-294  
C 40 18.2 46.7 601 3 US-09-949-016-65208  
C 41 18.2 46.7 601 3 US-09-949-016-83545  
C 42 18.2 46.7 601 3 US-09-949-016-83546  
C 43 18.2 46.7 601 3 US-09-949-016-83547  
C 44 18.2 46.7 601 3 US-09-949-016-83548  
C 45 18.2 46.7 653 3 US-09-533-559-6394

#### ALIGNMENTS

##### RESULT 1

US-09-270-767-13447  
; Sequence 13447, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 13447  
; LENGTH: 2182  
; TYPE: DNA  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-13447

Query Match 54.9%; Score 21.4; DB 3; Length 2182;  
Best Local Similarity 71.8%; Pred. No. 7.1;  
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATACGAGCGCTTATTCTAGACACTCACACGGCTCTT 39  
DB 1978 ATACGAGCGTATATATAGAAACATAACACGCATATT 2016

##### RESULT 2

US-09-949-016-15834/c  
; Sequence 15834, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15834

```
; LENGTH: 14641
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(14641)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15834

Query Match      52.3%; Score 20.4; DB 3; Length 14641;
Best Local Similarity 71.1%; Pred. No. 33;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      1 ATAGCGAGCGTCTTAATCTAGAACACTCACACGGCTCT 38
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      5467 ATAGAGAGTGTTTAATGCTAGAGACTAAGAGGCTACT 5430

RESULT 3
US-09-433-598-4/c
; Sequence 4, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
; APPLICANT: Taylor Dr., Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
; FILE REFERENCE: 07254/061001
; CURRENT APPLICATION NUMBER: US/09/433,598
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/107268
; EARLIER FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-433-598-4

Query Match      51.3%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATAGCGAGCGTCTTAATCTTA 20
      ||||| ||||| ||||| ||||| |||||
Db      20 ATAGCGAGCGTCTTAATCTTA 1

RESULT 4
US-09-848-838A-4/c
; Sequence 4, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)...(1036)
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated oligonucleotide
US-09-848-838A-4

Query Match      51.3%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATAGCGAGCGTCTTAATCTTA 20
      ||||| ||||| ||||| ||||| |||||
Db      20 ATAGCGAGCGTCTTAATCTTA 1

RESULT 5
US-09-433-598-1
; Sequence 1, Application US/09433598
; Patent No. 6238894
; GENERAL INFORMATION:
; APPLICANT: Taylor Dr., Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase
; FILE REFERENCE: 07254/061001
; CURRENT APPLICATION NUMBER: US/09/433,598
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: 60/107268
; EARLIER FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)..(1036)
US-09-433-598-1

Query Match      51.3%; Score 20; DB 3; Length 1119;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATAGCGAGCGTCTTAATCTTA 20
      ||||| ||||| ||||| ||||| |||||
Db      1100 ATAGCGAGCGTCTTAATCTTA 1119

RESULT 6
US-09-848-838A-1
; Sequence 1, Application US/09848838A
; Patent No. 6670160
; GENERAL INFORMATION:
; APPLICANT: Taylor, Diane
; APPLICANT: Wang, Ge
; APPLICANT: Palcic, Monica
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE
; FILE REFERENCE: 07254-061002
; CURRENT APPLICATION NUMBER: US/09/848,838A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 09/433,598
; PRIOR FILING DATE: 1999-11-02
; PRIOR APPLICATION NUMBER: US 60/107,268
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (137)...(1036)
```



US-09-848-838A-1

Query Match 51.3%; Score 20; DB 3; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATAGCGAGCGTCTAATTCTTA 20  
|||||  
Db 1100 ATAGCGAGCGTCTAATTCTTA 1119  
|||||

RESULT 7

US-09-513-999C-1023  
; Sequence 1023, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1023

; LENGTH: 386  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 87..386  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 383  
; OTHER INFORMATION: y=c or t  
US-09-513-999C-1023

Query Match 50.8%; Score 19.8; DB 3; Length 386;  
Best Local Similarity 69.2%; Pred. No. 25;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ATAGCGAGCGTCTAATTCTTAGAACACTCACACGCGTCTT 39  
|||||  
Db 310 ATAGGGATCGTCCAAACCGGAGACATACACCTCTCTTT 348  
|||||

RESULT 8

US-09-621-976-1260  
; Sequence 1260, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621.976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1260  
; LENGTH: 554  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 121..552  
; NAME/KEY: misc\_feature  
; LOCATION: 32

; OTHER INFORMATION: n=a, g, c or t  
US-09-621-976-1260

Query Match 50.8%; Score 19.8; DB 3; Length 554;  
Best Local Similarity 69.2%; Pred. No. 27;  
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ATAGCGAGCGTCTAATTCTTAGAACACTCACACGCGTCTT 39  
|||||  
Db 344 ATAGGGATCGTCCAAACCGGAGACATACACCTCTCTTT 382  
|||||

RESULT 9

US-09-902-540-316  
; Sequence 316, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902.540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 316

; LENGTH: 1807  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-316

Query Match 50.3%; Score 19.6; DB 3; Length 1807;  
Best Local Similarity 73.5%; Pred. No. 45;  
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 4 GCGAGCGTCTAATTCTTAGAACACTCACACGCGTC 37  
|||||  
Db 1261 GCGCCGTCGCATTTCTTAGACGAGCACACCGGTC 1294  
|||||

RESULT 10

US-09-433-598-3  
; Sequence 3, Application US/09433598  
; Patent No. 6238894  
; GENERAL INFORMATION:

; APPLICANT: Taylor Dr., Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: Alpha1, 2 Fucosyltransferase  
; FILE REFERENCE: 07254/061001  
; CURRENT APPLICATION NUMBER: US/09/433.598  
; CURRENT FILING DATE: 1998-11-02  
; EARLIER FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3

; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: oligonucleotide  
US-09-433-598-3

Query Match 48.7%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 GAACACTCACACGCGTCTT 39

Db 1 GAACACTCACACGGGTCTT 19  
|||||

## RESULT 11

US-09-848-838A-3  
; Sequence 3, Application US/09848838A  
; Patent No. 6670160  
; GENERAL INFORMATION:  
; APPLICANT: Taylor, Diane  
; APPLICANT: Wang, Ge  
; APPLICANT: Palcic, Monica  
; TITLE OF INVENTION: ALPHA 1,2-FUCOSYLTRANSFERASE  
; FILE REFERENCE: 07254-061002  
; CURRENT APPLICATION NUMBER: US/09/848,838A  
; CURRENT FILING DATE: 2002-05-21  
; PRIOR APPLICATION NUMBER: US 09/433,598  
; PRIOR FILING DATE: 1999-11-02  
; PRIOR APPLICATION NUMBER: US 60/107,268  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 19  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetically generated oligonucleotide  
US-09-848-838A-3

Query Match 48.7%; Score 19; DB 3; Length 19;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 GAACACTCACACGGGTCTT 39  
|||||  
Db 1 GAACACTCACACGGGTCTT 19

## RESULT 12

US-09-328-352-2920/c  
; Sequence 2920, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2920  
; LENGTH: 4515  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2920

Query Match 48.7%; Score 19; DB 3; Length 4515;  
Best Local Similarity 81.5%; Pred. No. 1.1e+02;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 GAGCGTCTAATTCCTAGAACACTCACAC 32  
|||||  
Db 4321 GAGCATCTAATGCAGACCACTCAAC 4295

## RESULT 13

US-08-894-344C-1  
; Sequence 1, Application US/08894344C  
; Patent No. 6172196  
; GENERAL INFORMATION:  
; APPLICANT: KAWASAKI, Hideki  
; APPLICANT: TOKAI, Masaya

; APPLICANT: KIRUCHI, Yasuhiro  
; APPLICANT: OUCHI, Kozo  
; TITLE OF INVENTION: DNA ENCODING PROTEIN COMPLEMENTING  
; TITLE OF INVENTION: YEAST  
; NUMBER OF INVENTION: LOW TEMPERATURE-SENSITIVE FERMENTABILITY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112-3801  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette - 3.50 inch, 1440 KB storage.  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS Ver3.30  
; SOFTWARE: PATENT AID Ver1.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/894,344C  
; FILING DATE: 15-AUGUST-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP343700/95  
; FILING DATE: 28-DECEMBER-1995  
; APPLICATION NUMBER: PCT/JP96/03862  
; FILING DATE: 27-DECEMBER-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perry, Lawrence S.  
; REGISTRATION NUMBER: 31865  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-218-2100  
; TELEFAX: 212-218-2200  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8874 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; ORIGINAL SOURCE:  
; ORGANISM: Saccharomyces cerevisiae  
; STRAIN: K2180-1B  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1 to 8874  
; IDENTIFICATION METHOD: E  
; FEATURE:  
; NAME/KEY: cleavage-site  
; LOCATION: 1291 to 1296  
; IDENTIFICATION METHOD: S  
; FEATURE:  
; NAME/KEY: cleavage-site  
; LOCATION: 4388 to 4393  
; IDENTIFICATION METHOD: S  
; FEATURE:  
; NAME/KEY: cleavage-site  
; LOCATION: 5927 to 5932  
; IDENTIFICATION METHOD: S  
; FEATURE:  
; NAME/KEY: cleavage-site  
; LOCATION: 7675 to 7680  
; IDENTIFICATION METHOD: S  
; US-08-894-344C-1

Query Match 48.7%; Score 19; DB 3; Length 8874;  
Best Local Similarity 81.5%; Pred. No. 1.3e+02;  
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 11 TCTAATTCCTAGAACACTCACACGGGTC 37  
|||||  
Db 7215 TTTAACTCTAGAACACACAAAAGGTC 7241

```

Query Match          48.7%; Score 19; DB 3; Length 8874;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels

QY      11  TCTAATTCTAGAACACTCACACGGGTC 37
      DB      7215  TTTAACTCTAGACACACAAAGGTC 7241

RESULT 15
US-09-125-028-1
; Sequence 1, Application US/09125028A
; Patent No. 6190707
; GENERAL INFORMATION:
; APPLICANT: WADOUX, Isabelle
; APPLICANT: COLAVIZZA, Didier
; APPLICANT: LOIEZ, Annie
; TITLE OF INVENTION: NOVEL COLD-SENSITIVE BREAD-MAKING YEASTS
; FILE REFERENCE: levure sensible froid
; CURRENT APPLICATION NUMBER: US/09/125,028A
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: PCT/FR97/00254
; EARLIER FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9621
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9318)
; OTHER INFORMATION: Use of n signifies any of g, a, c or t
US-09-125-028-1

Query Match          48.7%; Score 19; DB 3; Length 9621;
Best Local Similarity 81.5%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 5; Indels

QY      11  TCTAATTCTAGAACACTCACACGGGTC 37
      DB      7590  TTTAACTCTAGACACACAAAGGTC 7616

Search completed: January 30, 2006, 03:28:48
Job time : 75 secs

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 02:34:27 ; Search time 349 Seconds  
(without alignments)  
924.085 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3  
Perfect score: 39  
Sequence: 1 atagcggcgctctattctagacactcacgcgctctt 39

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 0.1

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgm2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgm2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
3: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
4: /cgm2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
5: /cgm2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
6: /cgm2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
7: /cgm2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
8: /cgm2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
9: /cgm2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
10: /cgm2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.4	57.4	653	3	US-09-988-067B-69
2	21	53.8	584	4	US-09-925-065A-336197
3	21	53.8	604	4	US-09-925-065A-336196
C 4	20.6	52.8	2409	6	US-10-369-493-26223
C 5	20.4	52.3	395	8	US-10-425-115-31553
C 6	20.4	52.3	488	4	US-09-925-065A-754744
7	20.2	51.8	451	5	US-10-198-846-375
8	20.2	51.8	686	5	US-10-198-846-1034
C 9	20.2	51.8	773	9	US-10-779-543-23653
C 10	20.2	51.8	2844	5	US-10-098-841-221
C 11	20.2	51.8	2871	8	US-10-788-792-120
C 12	20.2	51.8	2881	5	US-10-087-192-635
C 13	20.2	51.8	2881	7	US-10-058-270A-65
C 14	20.2	51.8	3230	5	US-10-198-846-10349
C 15	20.2	51.8	4604	5	US-10-087-192-634
C 16	20	51.3	20	3	US-09-848-838-4
C 17	20	51.3	20	7	US-10-663-033-4
C 18	20	51.3	577	4	US-09-925-065A-78749
C 19	20	51.3	934	3	US-09-895-913A-253
20	20	51.3	1119	3	US-09-848-838-1
21	20	51.3	1119	7	US-10-663-033-1
22	19.8	50.8	325	7	US-10-242-535A-10789
23	19.8	50.8	325	7	US-10-085-783A-10789

C 24	19.8	50.8	325	8	US-10-674-124A-11175	Sequence 11175, A
25	19.8	50.8	370	7	US-10-242-535A-22219	Sequence 22219, A
26	19.8	50.8	370	7	US-10-085-783A-22219	Sequence 22219, A
27	19.8	50.8	438	7	US-10-242-535A-34927	Sequence 34927, A
28	19.8	50.8	438	7	US-10-085-783A-34927	Sequence 34927, A
C 29	19.8	50.8	487	8	US-10-425-115-5839	Sequence 5839, Ap
30	19.8	50.8	488	7	US-10-242-535A-56833	Sequence 56833, A
31	19.8	50.8	488	7	US-10-085-783A-56833	Sequence 56833, A
32	19.8	50.8	569	8	US-10-357-930-35647	Sequence 35647, A
33	19.8	50.8	569	8	US-10-357-930-44454	Sequence 44454, A
34	19.8	50.8	570	8	US-10-357-930-14541	Sequence 14541, A
35	19.8	50.8	731	3	US-09-822-830A-72	Sequence 72, Appl
C 36	19.8	50.8	790	3	US-09-883-152-20	Sequence 20, Appl
C 37	19.8	50.8	800	3	US-09-986-480-32	Sequence 32, Appl
C 38	19.8	50.8	800	9	US-10-863-332-32	Sequence 32, Appl
40	19.8	50.8	826	5	US-10-027-632-145050	Sequence 145050, A
41	19.8	50.8	826	5	US-10-027-632-145051	Sequence 145051, A
42	19.8	50.8	826	6	US-10-027-632-145050	Sequence 145050, A
43	19.8	50.8	954	6	US-10-264-049-775	Sequence 775, App
44	19.8	50.8	1046	6	US-10-172-118-1642	Sequence 1642, Ap
45	19.8	50.8	1046	7	US-10-342-887-1642	Sequence 1642, Ap

## ALIGNMENTS

## RESULT 1

US-09-988-067B-69  
; Sequence 69, Application US/09988067B  
; Publication No. US200301241A1  
; GENERAL INFORMATION:  
; APPLICANT: Haas, Rainer  
; APPLICANT: Kleanthous, Harold  
; APPLICANT: Tomb, Jean-Francois  
; APPLICANT: Miller, Charles  
; APPLICANT: Al-Garawi, Amal  
; APPLICANT: Odenbreit, Stefan  
; APPLICANT: Meyer, Thomas  
; TITLE OF INVENTION: Helicobacter Polypeptides and  
; TITLE OF INVENTION: Corresponding Polynucleotide Molecules  
; FILE REFERENCE: 06132/040002  
; CURRENT APPLICATION NUMBER: US/09/988,067B  
; PRIOR FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: US 08/831,309  
; PRIOR FILING DATE: 1997-04-01  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69  
; LENGTH: 653  
; TYPE: DNA  
; ORGANISM: Helicobacter pylori  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (63)...(590)  
US-09-988-067B-69

Query Match 57.4%; Score 22.4; DB 3; Length 653;  
Best Local Similarity 95.8%; Pred. No. 14;  
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 16 TTCTAGAACACTCACACGCGTCTT 39  
|||||  
Db 547 TGCTAGAACACTCACACGCGTCTT 570

## RESULT 2

US-09-925-065A-336197  
; Sequence 336197, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single

```

: TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
: FILE REFERENCE: 108827.135
: CURRENT APPLICATION NUMBER: US/09/925,065A
: CURRENT FILING DATE: 2001-08-08
: PRIOR APPLICATION NUMBER: US 60/243,096
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: US 60/252,147
: PRIOR FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: US 60/250,092
: PRIOR FILING DATE: 2000-11-30
: PRIOR APPLICATION NUMBER: US 60/261,766
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/289,846
: PRIOR FILING DATE: 2001-05-09
: NUMBER OF SEQ ID NOS: 957086
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 336197
: LENGTH: 584
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-065A-336197

```

```

Query Match      53.8%; Score 21; DB 4; Length 584;
Best Local Similarity 75.0%; Pred. No. 56;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATACGAGCGGCTCTAATTTCTAGAACACTCACACGGCTC 37
    |||||
Db 545 ATATTTTGGAGTTTATTCTAGAAACTCACAAAGCCTC 581
    |||||

```

```

RESULT 3
US-09-925-065A-336196
; Sequence 336196, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of
; Nucleotide Polymorphisms in the Human
; HLA Class II Region
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 336196
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-336196

```

	Query Match	53.8%	Score 21	DB 4	Length 604
	Best Local Similarity	73.0%	Pred. No. 56		
	Matches 27	Conservative 0	Mismatches 10	Indels 0	Gaps 0
Qy	1	ATACGAGCGCTCTAAATCTAGAACACTCACACGGTC	37		
Db	469	ATATTTTGGATTTTATTTCTAGAAACTCACAGGCTC	505		

RESULT 4  
US-10-369-493-26223/c  
; Sequence 26223, Application US/10369493  
; Publication No. US20030233675A1

```

: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10 (52052) B
: CURRENT APPLICATION NUMBER: US/10/369,493
: CURRENT FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIOR FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 26223
: LENGTH: 2409
: TYPE: DNA
: ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26223

```

Query Match	52.8%	Score 20.6;	DB 6;	Length 2409;
Best Local Similarity	74.3%;	Pred. No. 1.2e+02;		
Matches 26;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;
Qy	4	GCAGCGCCTTAATTTCTAGAACACTC	CACACGCGTCT	38
Db	534	GAGAGCGCCTTAATTTCTAGGCAATAA	CACAGTTTAT	500

```

RESULT 5
US-10-425-115-31553/c
; Sequence 31553, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 31553
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(395)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_128786C.1
US-10-425-115-31553

```

```

Query Match      52.3%  Score 20.4;  DB 8;  Length 395;
Best Local Similarity 71.1%;  Pred. No. 95;
Matches 27;  Conservative 0;  Mismatches 11;  Indels 0;  Gaps 0;

Qy  1  ATAGCGAGCGCTCTAATTCTAGAAACACTTCACACGCGTCT 38
    |||||
Db   230 ATAGCGAGCGCTCTCAATCTCTACAGACATAGTTTCTTCT 193
    |||||

```

```

RESULT 6
US-09-325-065A-754744/c
; Sequence 754744, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

```

```
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75474
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
US-03-925-065A-754744

Query Match      52.3%; Score 20.4; DB 4; Length 488;
Best Local Similarity 71.1%; Pred. No. 1e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy      1 ATAGGAGCGTCTTAATTTCTAGAACACTCACAGCGTCT 38
Db      105 ATAGGAGATTCTATCTCTTAACACACACATCCGTAT 68

RESULT 7
US-10-198-846-375
; Sequence 375, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 446
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-375

Query Match      51.8%; Score 20.2; DB 5; Length 451;
Best Local Similarity 75.8%; Pred. No. 1.2e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTTAATTTCTAGAACACTCACAGC 34
Db      109 TTGGAGCCTCTATTCTTAGAACAGGAAACGC 141

RESULT 8
US-10-198-846-1034
; Sequence 1034, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1034
; LENGTH: 686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 632, 639, 657, 665, 673
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1034

Query Match      51.8%; Score 20.2; DB 5; Length 686;
Best Local Similarity 75.8%; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTTAATTTCTAGAACACTCACAGC 34
Db      220 TTGGAGCCTCTATTCTTAGAACAGGAAACGC 252

RESULT 9
US-10-779-543-23653/c
; Sequence 23653, Application US/10779543
; Publication No. US20050227917A1
; GENERAL INFORMATION:
; APPLICANT: Williams et al
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
; TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
; FILE REFERENCE: 2300-21302
; CURRENT APPLICATION NUMBER: US/10/779,543
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/076,555
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 09/217,471
; PRIOR FILING DATE: 1998-12-21
; PRIOR APPLICATION NUMBER: 60/068,755
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/080,664
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 09/297,648
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: PCT/US99/01619
; PRIOR FILING DATE: 1999-01-28
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 23767
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23653
; LENGTH: 773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 640..feature
; OTHER INFORMATION: n = A,T,C or G
US-10-779-543-23653

Query Match      51.8%; Score 20.2; DB 9; Length 773;
Best Local Similarity 75.8%; Pred. No. 1.4e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34
Db      151 TTGGAGCGCTCTATTCTTAGAACACGGAACGC 119

RESULT 10
US-10-098-841-221/c
; Sequence 221, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radje T.
; TITLE OF INVENTION: NO. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 221
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (355)..(2223)
US-10-098-841-221

Query Match      51.8%; Score 20.2; DB 5; Length 2844;
Best Local Similarity 75.8%; Pred. No. 1.8e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34
Db      2476 TTGGAGCGCTCTATTCTTAGAACACGGAACGC 2444

RESULT 11
US-10-098-841-221/c
; Sequence 221, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Drmanac, Radje T.
; TITLE OF INVENTION: NO. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 221
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (355)..(2223)
US-10-098-841-221

Query Match      51.8%; Score 20.2; DB 5; Length 2844;
Best Local Similarity 75.8%; Pred. No. 1.8e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34
Db      2476 TTGGAGCGCTCTATTCTTAGAACACGGAACGC 2444

RESULT 12
US-10-087-192-635/c
; Sequence 635, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 635
; LENGTH: 2881
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-635

Query Match      51.8%; Score 20.2; DB 5; Length 2881;
Best Local Similarity 75.8%; Pred. No. 1.8e+02;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34
Db      2122 TTGGAGCGCTCTATTCTTAGAACACGGAACGC 2090

RESULT 13
US-10-058-270A-65/c
; Sequence 65, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afar, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
```



; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer

; FILE REFERENCE: 018501-005210US  
; CURRENT APPLICATION NUMBER: US/10/058,270A  
; CURRENT FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: US 60/263,965  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: US 60/265,928  
; PRIOR FILING DATE: 2001-02-02  
; PRIOR APPLICATION NUMBER: US 09/829,472  
; PRIOR FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 60/282,698  
; PRIOR FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 60/288,590  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/294,443  
; PRIOR FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 141  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 65  
; LENGTH: 2881  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-058-270A-65

Query Match 51.8%; Score 20.2; DB 7; Length 2881;  
Best Local Similarity 75.8%; Pred. No. 1.8e+02;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34  
Db 2122 TTGGAGCCTCTATTCTTAGAACAGGAACGC 2090

## RESULT 14

US-10-198-846-10349/c  
; Sequence 10349, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 10349  
; LENGTH: 3230  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 3212, 3213, 3214, 3215, 3216, 3217, 3218, 3219, 3220, 3221,  
; LOCATION: 3222, 3223, 3224, 3225, 3226, 3227, 3228, 3229, 3230  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-10349

Query Match 51.8%; Score 20.2; DB 5; Length 3230;  
Best Local Similarity 75.8%; Pred. No. 1.9e+02;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34  
Db 2471 TTGGAGCCTCTATTCTTAGAACAGGAACGC 2439

## RESULT 15

US-10-087-192-634/c  
; Sequence 634, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: PastSeq for Windows Version 4.0  
; SEQ ID NO 634  
; LENGTH: 46404  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(46404)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-634

Query Match 51.8%; Score 20.2; DB 5; Length 46404;  
Best Local Similarity 75.8%; Pred. No. 3.4e+02;  
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 TAGCGAGCGTCTAATTTCTAGAACACTCACACGC 34  
Db 35645 TTGGAGCCTCTATTCTTAGAACAGGAACGC 35613

Search completed: January 30, 2006, 05:23:22  
Job time : 350 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2006, 03:26:26 ; Search time 392.5 Seconds  
(without alignments)  
82.538 Million cell updates/sec

Title: US-10-663-033-REVCOMP4-SEQ3

Perfect score: 39  
Sequence: 1 ataggcgctctattcttagaacactcacgcgtctt 39

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.1

Searched: 6059551 seqs, 415333918 residues

Total number of hits satisfying chosen parameters: 12119102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.New.\*

- 1: /cgm2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 2: /cgm2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /cgm2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /cgm2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 5: /cgm2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 6: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 8: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 9: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 10: /cgm2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 11: /cgm2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.2	51.8	3342	7	US-10-750-185-34911
2	20.2	51.8	3342	7	US-10-750-623-34911
3	19.6	50.3	385	8	US-11-128-061-2345
4	19.6	50.3	385	8	US-11-128-061-5987
5	19.6	50.3	385	8	US-11-128-049-2345
6	19.6	50.3	385	8	US-11-128-049-5987
7	19.6	50.3	142605	8	US-11-121-086-64
8	19.2	49.2	127917	7	US-10-775-169-82
9	19	48.7	1419	7	US-10-750-185-46559
10	19	48.7	1419	7	US-10-750-623-46559
11	18.8	48.2	201	7	US-10-995-561-64263
12	18.8	48.2	201	7	US-10-995-561-64296
13	18.8	48.2	56448	7	US-10-995-561-13369
14	18.6	47.7	930	8	US-11-055-822-509
15	18.6	47.7	169725	8	US-11-121-086-63
16	18.4	47.2	1981	7	US-10-750-185-58250
17	18.4	47.2	1981	7	US-10-750-623-58250
18	18.2	46.7	1124	7	US-10-750-185-24993
19	18.2	46.7	1124	7	US-10-750-623-24993
20	18.2	46.7	1836	7	US-10-750-185-57119
21	18.2	46.7	1836	7	US-10-750-623-57119
22	18.2	46.7	3581	7	US-10-750-185-35045

ALIGNMENTS

RESULT 1

US-10-750-185-34911

; Sequence 34911, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

; APPLICANT: BATES, Stephen

; APPLICANT: FANTIN, Dennis

; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10750185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: Patent version 3.1

; SEQ ID NO 34911

; LENGTH: 3342

; TYPE: DNA

; ORGANISM: Bovine 19866880569723

US-10-750-185-34911

Query Match 51.8%; Score 20.2; DB 7; Length 3342;

Best Local Similarity 75.8%; Pred. No. 9.7;

Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4 GCGAGCGCTCTAATCTTAGAACACTCACACGGT 36

Db 2328 CGGCTTTAGAAATCTTGGAACTCACACGCT 2360

RESULT 2

US-10-750-623-34911

; Sequence 34911, Application US/10750623

; Publication No. US20050287531A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.

; APPLICANT: KERR, Richard

; APPLICANT: ROSENFELD, David

; APPLICANT: HOLM, Tom

```
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34911
; LENGTH: 3342
; TYPE: DNA
; ORGANISM: Bovine 19866880569723
US-10-750-623-34911

Query Match      51.8%; Score 20.2; DB 7; Length 3342;
Best Local Similarity 75.8%; Pred. No. 9.7;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      4 GCGAGCGTCTAATTCTAGAACACTCACACGCGT 36
Db      2328 GCGTCTTAGAATTCTGGACACTCACAAAGCGT 2360

RESULT 3
US-11-128-061-2345
; Sequence 2345, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2345
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-2345

Query Match      50.3%; Score 19.6; DB 8; Length 385;
Best Local Similarity 73.5%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
Db      105 GAGAAATTTCAATTCCTGGGACACTCACATGAGTCCT 138

RESULT 4
US-11-128-061-5987
; Sequence 5987, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
```

```
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5987
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-5987

Query Match      50.3%; Score 19.6; DB 8; Length 385;
Best Local Similarity 73.5%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
Db      105 GAGAAATTTCAATTCCTGGGACACTCACATGAGTCCT 138

RESULT 5
US-11-128-049-2345
; Sequence 2345, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2345
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-2345

Query Match      50.3%; Score 19.6; DB 8; Length 385;
Best Local Similarity 73.5%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      6 GAGCGTCTAATTCTAGAACACTCACACGCGTCTT 39
Db      105 GAGAAATTTCAATTCCTGGGACACTCACATGAGTCCT 138

RESULT 6
US-11-128-049-5987
; Sequence 5987, Application US/11128049
; Publication No. US20060010513A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
```

```
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR
; TITLE OF INVENTION: MAKING AND USING SAME
; FILE REFERENCE: 01997.027700
; CURRENT APPLICATION NUMBER: US/11/128,049
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5987
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-049-5987

Query Match      50.3%; Score 19.6; DB 8; Length 385;
Best Local Similarity 73.5%; Pred. No. 12;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy      6 GAGCGTCTAAATTTCTAGAACACTCACCACGCGTCT 39
Db      105 GAGAAATTCATCTCGGACACTCAGATGCTCT 138

RESULT 7
US-11-121-086-64
; Sequence 64, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 64
; LENGTH: 142605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-64

Query Match      50.3%; Score 19.6; DB 8; Length 142605;
Best Local Similarity 84.6%; Pred. No. 41;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      7 AGCGTCTAAATTTCTAGAACACTCAGAC 32
Db      117976 AGCGTCTAAATTTAGAACTCTTGAC 118001

RESULT 8
US-10-775-169-82
; Sequence 82, Application US/10775169
; Publication No. US20050287532A9
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dorner, Andrew
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 82
; LENGTH: 127917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-82

Query Match      49.2%; Score 19.2; DB 7; Length 127917;
Best Local Similarity 87.5%; Pred. No. 62;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      15 ATTCTAGAACACTCACCACGCGTCT 38
Db      40267 ATGCTAGACACTCACCACGCGTCT 40290

RESULT 9
US-10-750-185-46559/c
; Sequence 46559, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46559
; LENGTH: 1419
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-46559

Query Match      48.7%; Score 19; DB 7; Length 1419;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy      4 GCAGCGCTCTAAATTTCTAGAACACTCACCACGCGTCT 38
Db      597 GCATTCATATATTTCTTAAGCACACATGCACTT 563

RESULT 10
US-10-750-623-46559/c
; Sequence 46559, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46559
; LENGTH: 1419
; TYPE: DNA
```



; LOCATION: (101)..(907)  
 ; OTHER INFORMATION: RXC02080  
 US-11-055-822-509

Query Match 47.7%; Score 18.6; DB 8; Length 930;  
Best Local Similarity 84.0%; Pred. NO. 41;  
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 AGCGAGCGTCTAATTTCTAGAACACT 27  
Db 12 AGCGAGCTTCTAATTTCTAGCAAGCT 36

## RESULT 15

```

US-11-121-086-63/c
; Sequence 63, Application US/11121086
; Publication NO. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138-6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

```

Query Match 47.7%; Score 18.6; DB 8; Length 169725;  
Best Local Similarity 72.7%; Pred. NO. 1.2e+02;  
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 7 AGCGTCTAATTTCTAGAACACTCACAGCGTC TT 39  
||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 89829 AGGCTCTTTTTCAAGAAGACTCACCTCTTA TT 89797

Search completed: January 30, 2006, 05:36:33  
Job time : 393.5 secs

**This Page Blank (uspto)**